NBER WORKING PAPER SERIES

THE SKILLS TO PAY THE BILLS: RETURNS TO ON-THE-JOB SOFT SKILLS TRAINING

Achyuta Adhvaryu
Namrata Kala
Anant Nyshadham

Working Paper 24313
http://www.nber.org/papers/w24313

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2018

We are very grateful to Anant Ahuja, Chitra Ramdas, Raghuram Nayaka, Sudhakar Bheemarao, and Paul Ouseph for their coordination, enthusiasm, and guidance. Thanks to Dotti Hatcher, Lucien Chan, Noel Simpkin, and others at Gap, Inc. for their support and feedback on this work. We acknowledge funding from Private Enterprise Development in Low-Income Countries (PEDL) initiative, and Adhvaryu's NIH/NICHD (5K01HD071949) career development award. This research has benefited from discussions with Michael Boozer, Robert Gibbons, Paul Gertler, Markus Goldstein, Rocco Macchiavello, David McKenzie, Dilip Mookherjee, Claudia Olivetti, Antoinette Schoar, Tavneet Suri, Chris Udry, John Van Reenen, and Chris Woodruff, and seminar audiences at NBER, Penn, MIT, USC, Madrid, PEDL, IGC, World Bank, AEA, Northeastern, NEUDC, Chicago, Michigan, McGill, Georgetown, and Copenhagen. Many thanks to Lavanya Garg, Robert Fletcher, and Aakash Mohpal for excellent research assistance. The views expressed herein do not represent PEDL, NIH, Gap, Inc., or Shahi Exports. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Skills to Pay the Bills: Returns to On-the-job Soft Skills Training
Achyuta Adhvaryu, Namrata Kala, and Anant Nyshadham
NBER Working Paper No. 24313
February 2018
JEL No. J24,M53,O15

ABSTRACT

We evaluate the causal impacts of on-the-job soft skills training on the productivity, wages, and retention of female garment workers in India. The program increased women's extraversion and communication, and spurred technical skill upgrading. Treated workers were 20 percent more productive than controls post-program. Wages rise very modestly with treatment (by 0.5 percent), with no differential turnover, suggesting that although soft skills raise workers' marginal products, labor market frictions are large enough to create a substantial wedge between productivity and wages. Consistent with this, the net return to the firm was large: 258 percent eight months after program completion.

Achyuta Adhvaryu
Ross School of Business
University of Michigan
701 Tappan Street
Ann Arbor, MI 48109
and NBER
adhvaryu@umich.edu

Namrata Kala
Harvard University
27 Hillhouse Ave
New Haven, CT 06511
namrata.kala@gmail.com

Anant Nyshadham
Department of Economics
Boston College
Maloney Hall, 324
Chestnut Hill, MA 02467
and NBER
nyshadha@bc.edu

An online appendix is available at http://www.nber.org/data-appendix/w24313

1 Introduction

Soft skills – e.g., teamwork, leadership, relationship management, personality factors, effective time allocation, and the ability to assimilate information – are highly predictive of success in the labor market (Bassi et al., 2017; Borghans et al., 2008; Deming, 2015; Groh et al., 2015; Guerra et al., 2014; Heckman and Kautz, 2012; Heckman et al., 2006; Montalvao et al., 2017). Surveys of employers from around the world corroborate that soft skills are in great demand, and that firms often struggle to find workers with high levels of these skills (Cunningham and Villaseñor, 2016).

Studies from psychology and economics demonstrate that it is possible to inculcate soft skills in early childhood, via, for example, home-based stimulation and high quality preschool programs (Attanasio et al., 2014; Gertler et al., 2014; Grantham-McGregor et al., 1991; Ibarrarán et al., 2015). But how malleable soft skills are in adulthood, and whether training programs that aim to increase the stock of these skills can indeed generate causal impacts on productivity, have only begun to be explored (Acevedo et al., 2017; Ashraf et al., 2017; Campos et al., 2017; Groh et al., 2012). It is not obvious that inculcating these skills in a meaningful way is possible: structural estimates of dynamic human capital accumulation models suggest that it may indeed be difficult to affect non-cognitive skill levels at later ages, particularly for those with low baseline stocks, due to dynamic complementarities (Aizer and Cunha, 2012; Cunha et al., 2010; Heckman and Mosso, 2014).

Moreover, when general training is delivered within the firm (as it often is[1]), it is imperative to know the firm's returns to training in addition to worker productivity effects. This impact, in turn, is governed by labor market structure. In perfectly competitive markets, workers' wages would need to increase commensurate to their marginal products; any firm that paid below marginal product would lose the newly trained workers as they received higher wage offers at other firms. As Becker (1964) famously noted, this implies that with perfect labor markets, even general training programs that generate large productivity returns may not be appealing investments for firms. On the other hand, if asymmetric information and search frictions play a role in the labor market, then the resulting wedge between workers' marginal products and their wages in equilibrium may create positive productivity rents from general training for firms (Acemoglu, 1997; Acemoglu and Pischke, 1998, 1999; Autor, 2001; Chang and Wang, 1996; Katz and Ziderman, 1990). Since most soft skills are "general," the extent of labor market frictions thus likely polices the ability to deliver soft skills training through firms, even when training raises productivity.

The questions that motivate our study, then, are threefold. First, is it possible to improve soft skills meaningfully for workers with low stocks of these skills? Second, if skills do improve, what are the causal impacts on workplace outcomes, including productivity, wages, and retention? Finally does it pay for firms to provide on-the-job soft skills training to workers, and what does this rate of return tell us about the nature of labor market frictions as pertains to soft skills?

To answer these questions, we partnered with the largest ready-made garment export firm in India to evaluate an intensive, workplace-based soft skills training program. The initiative, which is named Personal Advancement and Career Enhancement (P.A.C.E.), aims to empower female garment workers via training in a broad variety of life skills, including modules on communication, time management,

[1]See, e.g., Bassanini et al. (2007).

financial literacy, successful task execution, and problem-solving. These skills are important inputs into production in the ready-made garments context. Workers need effective communication to resolve throughput issues with other team members (e.g., identifying and working through bottlenecks in real time). They need relationship management skills to relay information in a productive way to supervisors (e.g., machine malfunction, requesting breaks or help to complete tasks, etc.). And they need problem-solving frameworks to effectively deal with daily shocks to production.

We conducted a randomized controlled trial (RCT) in five garment factories in urban Bengaluru, India. We assessed the impacts of soft skills training on 1) direct and indirect measures of the stock of these skills; and 2) administrative data on retention, productivity, wages, task complexity, and other workplace outcomes. Finally, we compute the firm's returns, combining our point estimates with data on the program's costs and the firm's accounting profits.

We enrolled female garment workers in a lottery for the chance to take part in the P.A.C.E. program and used a two-stage randomization procedure to assign workers to treatment. In the first stage, we randomized production lines to treatment. In the second stage, within treatment lines, we randomized workers who had enrolled in the lottery to either direct P.A.C.E. training or spillover treatment. We thus estimate treatment effects by comparing trained workers (on treatment lines) to control workers on control lines (who enrolled in the lottery but whose lines were assigned to control). We estimate spillovers by comparing untrained workers on treatment lines to control workers on control lines.

Endline survey results for treated and control workers and pre/post-module testing of treated workers indicate that stocks of soft skills improved in several important dimensions. Specifically, treated women showed a pronounced increase in extraversion, which may impact productivity via improvements in the ability to communicate and solve issues collaboratively with peers and supervisors. These women were also more likely to request and complete technical skill development trainings, generating complementary improvements in "hard" skills. Survey results indicate greater self-assessment of workplace quality (relative to peers of the same technical skill grade), consistent with an increase in self-regard. Finally, pre/post data from assessment tools designed to measure learning in each of the program's modules show that initial stocks of knowledge in each of the program's target areas were low, and that treated workers substantially improved these stocks through the program (most markedly for communication skills).

Direct impacts on workplace outcomes, measured using the firm's administrative data, are consistent with the acquisition of soft skills by workers. Treated workers are more productive by about 11 percentage points (20% higher than the control mean) and more likely to be assigned to complex tasks. Impacts last up to 8 months after program completion (when we ceased data collection), suggesting that learned skills translated into persistent improvements in workplace outcomes. Workers on treatment lines who did not receive the program are also more productive and are assigned to more complex operations, generating team-level (production line) impacts on productivity post-program completion. Wages went up very slightly as a result of treatment: an increase of about 0.5 percent. The program had no sustained impact on turnover. Retention was actually higher in the treatment group relative to control during the program period; this effect diminished after program completion.[2]

[2]We use a dynamic inverse probability weighting procedure, described in detail in section 4, throughout our analysis to correct for potential changes in the size and composition of the treatment and control groups over time.

Taken in sum, we interpret the results to indicate that the program increased workers' stocks of soft skills, which in turn led to productivity improvements.[3] Combined with the fact that there was essentially no impact on wage or long-run turnover, our results suggest the presence of substantial labor market frictions that prevent workers from capturing more of the productivity rents that ensue from training (Acemoglu, 1997; Acemoglu and Pischke, 1999). The nature of the hiring process in this labor market helps to rationalize this result. Specifically, sewing machine operators are evaluated – and accordingly are given wage offers – based only on stitching skills. Soft skills are largely unobserved in this hiring process and therefore are not priced into the wage, in line with other hiring processes for frontline workers in low-income country contexts (Bassi et al., 2017). This information friction likely generates the observed difference in impacts of soft skills training on marginal productivity as compared to wage.

We use our estimates of impacts on workplace outcomes along with program cost and accounting profit data to calculate the costs and benefits of the program to the firm. The net rate of return was 73% by the end of the program period. Eight months after program completion, fueled by post-program increases in productivity, the return climbed to over 250%. These large returns are rationalized by the relatively low costs of the program combined with the accumulated effects on productivity and person days, and are consistent with other recent interventions in garment factories in South Asia (Menzel, 2015).

Our main contribution is to the study of soft skills in the labor market. We join a handful of recent studies that evaluate the causal impacts of soft skills training on economic outcomes (Acevedo et al., 2017; Ashraf et al., 2017; Campos et al., 2017; Groh et al., 2012; Schoar, 2014). We add to this work by studying training within the firm, which emphasizes estimating firms' returns, tying our work to the literature on the role of labor market frictions in firms' decisions to train their workers (Acemoglu, 1997; Acemoglu and Pischke, 1998, 1999; Autor, 2001). We are also able to directly estimate impacts on individual productivity, which is missing from previous work.[4]

Other previous work quantifying the productivity impacts of on-the-job training generally uses observational data on firms and workers in the United States and Western Europe (Barrett and O'Connell, 2001; Barron et al., 1999; Dearden et al., 2006; Konings and Vanormelingen, 2015; Mincer, 1962). These studies tend to find that training increases productivity, but there is disagreement on the magnitude of this increase (Blundell et al., 1999). Specifically, when endogeneity of training is accounted for (e.g., using matching methods), productivity returns become quite small (Goux and Maurin, 2000; Leuven and Oosterbeek, 2008). We add to this literature in three ways. First, we estimate causal effects by exploiting randomized assignment to training, which overcomes potential self-selection bias (Altonji and Spletzer, 1991; Bartel and Sicherman, 1998). Second, we estimate impacts on retention in addition to productivity; retention is crucial to understanding firms' overall returns to training but has not been examined thus far. Third, we carry out our experiment in a low-income country setting, where training frontline workers might have large potential given low levels of baseline skills.

The rest of the paper is organized as follows. Section 2 discusses the garment production context

[3] We address several other possible mechanisms in section 6, including potential changes to mental and physical health, reciprocity, and social capital.

[4] Campos et al. (2017) measure microenterprise profits, which of course are in part a function of productivity.

and reviews the details of the training program and the experimental design. Section 3 discusses the data sources and the construction of key variables, and section 4 describes the estimation strategy. Section 5 describes the results of the estimation. Section 6 discusses and evaluates possible mechanisms and presents an analysis of the costs and benefits to the firm. Section 7 concludes.

2 Context, Program Details, and Experiment Design

2.1 Context

2.1.1 Ready-made Garments in India

Apparel is one of the largest export sectors in the world, and vitally important for the economies of several large developing countries (Staritz, 2010). India is one of the world's largest producers of textile and garments, with export value totaling $10.7 billion in 2009-2010. The size of the sector and the labor-intensity of the garment production process make the sector well-suited to absorb the influx of young, unskilled and semi-skilled labor migrating from rural self-employment to wage labor in urban areas, especially women (World Bank, 2012). Women comprise the majority of the workforce in garment factories, and new labor force entrants tend to be disproportionately female, particularly in countries like India where the baseline female labor force participation rate is low (Staritz, 2010). Shahi Exports, Private Limited, the firm with which we partnered to do this study, is the largest private garment exporter in India, and the single largest private employer of unskilled and semi-skilled female labor in the country.

2.1.2 The Garment Production Process

There are three broad stages of garment production: cutting, sewing, and finishing. In this study, we estimate program impacts on workers from the sewing department only, as measures of individual productivity and task complexity are only available for sewing workers.[5] Sewing department workers make up about 80% of the factory's total employment.

In the sewing department of the study factories (as in most medium and large garment factories), garments are sewn in production lines consisting of around 50-70 workers arranged in sequence. Most of the workers on the line are assigned to machines completing sewing tasks (one person to a machine). The remaining workers perform complementary tasks to sewing, such as folding or aligning the garment to feed it into a machine. Each line produces a single style of garment at a time.[6]

The line is subdivided into smaller groups of operations that produce subsections of the garment (e.g., collar, sleeve, or pocket). These groups are separated by "feeding points" at which the prepared materials for each subsection of the garment are fed in bundles (e.g., materials for 20 pockets or collars of the current shirt will be fed at one point and materials for 40 sleeves will be fed at the next point). This structure of subdivisions, multiple feeding points, and bundles of materials is very common in

[5]This is because a standardized measure of output is recorded for each worker in each hour on the sewing floor, but such a measure is not recorded for workers in other departments.

[6]The color and size of the garment might vary but the design and style will be the same for every garment produced by that line until the ordered quantity for that garment is met.

the industry (and in fact mirrored in many other manufacturing industries) and is used explicitly to decouple, as much as is possible, productivity at adjacent operations or subdivisions and allow time for rebalancing of productivity across the line.

Completed sections of garments pass between machine operators in these bundles, are attached to each other in additional operations along the way, and emerge at the end of the line as completed garments. These completed garments are then transferred to the finishing floor. In the finishing department, garments are checked, ironed, and packed for shipping. Most quality checking is done on the sewing floor during production, but final checks are done in the finishing stage. Any garments with quality issues are sent back to the sewing floor for rework or, if irreparably ruined, are discarded before packing.[7] Orders are then packed and sent to ports for export.

2.2 Program Details

The Personal Advancement and Career Enhancement (P.A.C.E.) program was designed and first implemented by Gap, Inc. for female garment workers in low-income contexts. Shahi Exports participated in the original design and piloting of the program as one of the largest suppliers to Gap. The intervention we study involved the implementation of the P.A.C.E. program in five factories in the Bengaluru area which had not yet adopted the program. The goal of this 80-hour program was to improve life skills such as time management, effective communication, problem-solving, and financial literacy for its trainees. The program began with an introductory ceremony for participants, trainers, and firm management. The core modules were: Communication (9.5 hours); Problem Solving and Decision-Making (13 hours); Time and Stress Management (12 hours); Execution Excellence (5 hours); Financial Literacy (4.5 hours); and Legal Literacy and Social Entitlements (8.5 hours).[8] Table A1 provides an overview of the topics covered in each module. After all modules had been completed, there were two review sessions (3 hours in total) reiterating concepts from early modules and discussing how participants would apply their learning to personal and professional situations. At the close of the program there was a graduation ceremony.

Workers participated in two hours of training per week. Management allocated one hour of workers' production time a week to the program, and workers contributed one hour of their own time. Training sessions were conducted at the beginning of the production day in designated classroom spaces in the factories, with workers assigned to groups corresponding to different days of the work week. That is, a worker assigned to the Monday group would be expected to attend training starting one hour before production starts on each Monday and ending after the first production hour of the day is completed (two hours in total). Production constraints required that each day's group be composed of workers from across production lines so as not to produce large, unbalanced absences from any one line in the first hour of any production day. Accordingly, the training groups were balanced in size with roughly 50 trainees per class and no more than 3-4 from a given line in each group.

[7] Completed quantities of garments recorded in the production data reflect only pieces which have passed quality checks, so quantity produced reflects both quantity and minimum quality combined.

[8] Additional modules on Water, Sanitation and Hygiene (6 hours) and General and Reproductive Health (10 hours) were also included, but were not considered core modules. Pre/post assessments were not conducted for ancillary modules such as sanitation.

Due to holidays and festivals (which are times of high absenteeism), sessions were conducted in practice somewhat more flexibly with respect to timing. Catch-up sessions were conducted for workers who were unable to attend a session. This flexibility is reflected in average attendance (of non-attrited workers) to the core program modules, which was very high, ranging between 94 and 99 percent (see Figure A12). With these adjustments, overall program implementation took about 12 months: the introductory ceremony was in July 2013, training was conducted between July 2013 and June 2014, and the closing ceremony in July 2014.

Figure 1A: Experimental Design

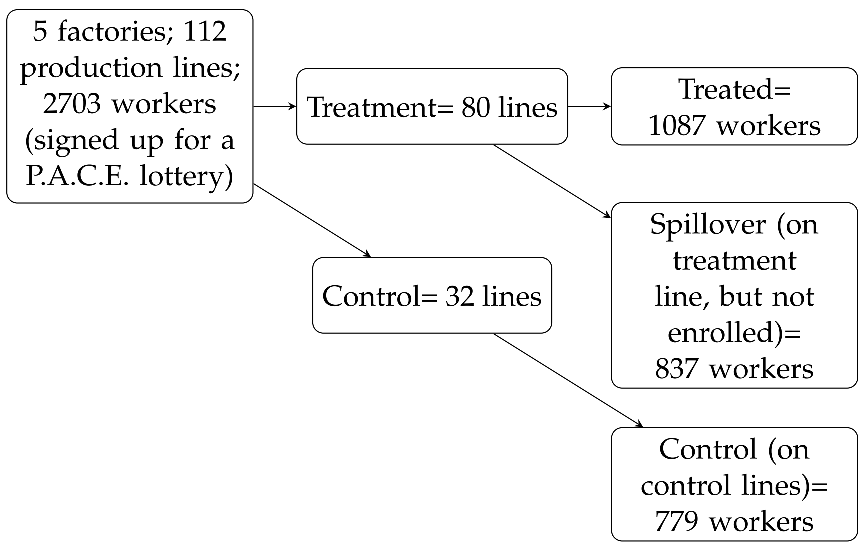


Figure 1B: Timeline of Experiment and Data Collection

January 2013 • Salary and Attendance Data Collection Starts
June 2013 • Treatment Assignment Announcement and Productivity Data Collection Starts
July 2013 • Training Program Starts (Pre and Post Module Testing During Training)
June 2014 • Training Program Ends and Worker Survey Conducted
December 2014 • Salary Data Collection Ends
February 2015 • Attendance and Production Data Collection Ends

2.3 Experimental Design

Participants were chosen from a pool of workers who expressed interest and committed to enroll in the program. The workers were informed that the training was oversubscribed and that a subset of workers would be chosen at random from a lottery to actually receive the training, with untreated

Figure 1C: Data Type and Availability

Attendance & Late-coming	Daily (January 2013-February 2015)
Productivity	Daily (June 2013-February 2015)
Salary	Monthly (January 2013-December 2014)
Survey Outcomes	Cross-sectional (June 2014)
Retention	Daily from Productivity Data, Monthly from Salary Data

workers granted the right to enroll in a later lottery for the next training batch.[9] Randomization was conducted at two levels: line level (stratified by factory unit, above- and below-median baseline efficiency and above- and below-median baseline attendance, and above- and below-median enrollment in the lottery), and then at the individual level within treatment lines. The five factory units had 112 production lines in total. In the first stage of randomization, roughly two-thirds of production lines within each factory unit were randomized to treatment, yielding 80 treatment lines and 32 control lines across units. In the second stage of randomization, within lines randomized to treatment, a fixed number of workers (13-14) from each treatment line were randomly chosen to take part in the P.A.C.E. program from the total set of workers who expressed interest by enrolling in the treatment lottery.[10]

Figure 1A presents a schematic diagram of the experimental design. 2703 workers signed up for the treatment lottery, from which 1087 were chosen for treatment. Out of the 1616 untrained workers, 779 workers were in control lines, and the remainder, 837 workers, were in treatment lines. The former group (untrained workers in control lines) serves as our primary control. The latter group (untrained workers in treatment lines) is used to estimate treatment spillovers. Summary statistics and balance checks are discussed in Section 3.4.[11] Figure 1B presents the timeline of the experiment and data collection.

3 Data

Figure 1C presents an overview of the different data sources used in the evaluation, the frequency of data collection of each data type, and the availability of the data over time. Details of the variables of interest are presented below.

3.1 Production Data

Productivity data were collected using tablet computers assigned to each production line on the sewing floor. The employee in charge of collecting the data (the "production writer"), who was prior to our

[9]Importantly, losers of the lottery were told that they would not necessarily receive the training in the next batch, nor would they be able to earn the right to be trained in any way, but rather that subsequent training batches would also be chosen at random via lottery.

[10]The decision to allocate a fixed number of workers to treatment per treatment line was due primarily to production constraints requiring a minimum manpower be present at all times during production hours.

[11]For the sake of brevity, we present only balance checks for treatment versus control workers, but balance holds across spillover versus control workers as well (results available upon request).

intervention charged with recording by hand on paper each machine operator's completed operations each hour for the line, was trained to input production data directly in the tablet computer instead. These data then automatically wirelessly synced to the server. Importantly, from the perspective of the garment workers, production data were being recorded identically before, during, and after the intervention across treatment and control lines. Note that though productivity was being recorded prior to the program implementation, the worker-hourly level data was not kept prior to the introduction of the tablet computers for production writing but rather discarded after line-daily level aggregate measures were input into the data server. Accordingly, line-daily level aggregate data was all that was available at the time of treatment assignment, and as mentioned above, the first stage randomization of lines to treatment was stratified by line-level baseline efficiency.

3.1.1 Productivity

The key measure of productivity we study is efficiency. Efficiency is calculated as pieces produced divided by the target quantity of pieces per unit time. In order to calculate the worker-level daily mean of production from these observations, we average the efficiency of each worker over the course of the day (8 production hours).[12]

At the worker-hour level, we define pieces produced as the number of garments that passed a worker's station by the end of that production hour. For example, if a worker was assigned to sew plackets onto shirt fronts, the number of shirt fronts at that worker's station that had completed placket attachment by the end of a given production hour would be recorded as that worker's "pieces produced." The target quantity for a given operation is calculated using a measure of garment and operation complexity called the "standard allowable minute" (SAM). SAM is defined as the number of minutes required for a single garment of a particular style to be produced. That is, a garment style with a SAM of 30 is deemed to take half an hour to produce one complete garment. This measure at the line level is then decomposed into worker or task specific increments. A line with 60 machine operators then would have an average worker-hourly SAM of 0.5.[13] As the name suggests, it is standardized across the global garment industry and is drawn from an industrial engineering database.[14] The target quantity for a given unit of time for a worker completing a particular operation is then calculated as the unit of time in minutes divided by the SAM. That is, the target quantity of pieces to be produced by a worker in an hour for an operation with a SAM of 0.5 will be $60/.5 = 120$.

As mentioned in the previous section, hourly productivity data was available starting the month of treatment announcement. During the month of treatment announcement (June 2013) the tablets were introduced onto the production floors. Accordingly, June 2013 represents the pre-program baseline for all productivity analysis below.

[12]As noted above, pieces are recorded only if the garment is complete and passes minimum quality standards during in-line and end-line quality checking. In averaging across hourly quantities within the day, we expect that mis-measurement arising from re-worked defective pieces is minimized.

[13]Mean SAM across worker hourly observations is 0.61 with a standard deviation of 0.20.

[14]This measure may be amended to account for stylistic variations from the representative garment style in the database. Any amendments are explored and suggested by the sampling department, in which master tailors make samples of each specific style to be produced by lines on the sewing floor (for costing purposes).

3.2 Human Resources Data: Attendance and Salary

Data on demographic characteristics, attendance, tenure and salary of workers are kept in a firm-managed database. The data linked to worker ID numbers were shared with us. The variables available in demographic data include age, date on which the worker joined the firm, gender, native language, and education. We combined these with daily attendance data at the worker level indexed by worker ID number and date, which records whether a worker attended work on a given date, whether absence was authorized or not, and whether a worker was late to work on a given day (worker tardiness). We also combined these with monthly salary data which also indicates current skill grade level. The salary data are available until six months post-program completion, unlike the productivity and attendance data, which are available for eight months after program completion.

3.3 Survey Data

In addition to measuring workplace outcomes, a survey of 1000 randomly chosen treated and control workers was conducted in June 2014, the month of program completion. The survey covered, among other things, questions related to financial decisions (including savings and debt) and awareness of and participation in welfare programs (government or employer sponsored). It also measured personality characteristics (conscientiousness, extraversion, locus of control, perseverance, and self-sufficiency), mental health (hope/optimism, self-esteem, and the Kessler 10 module, which can be used to diagnose moderate to severe psychological distress (Kessler et al., 2003)), and risk and time preferences elicited using lottery choices.[15] Finally, the survey covered worker's self-assessments relative to peers by asking them to imagine a six-step ladder with the lowest productivity workers on the lowest steps, and then asking them which step they would place themselves on; participation in skill development programs; production awards; and incentive programs on the job.

3.4 Summary Statistics and Balance Checks

Table 1 presents summary statistics of the main variables of interest, as well as balance checks for baseline values of attendance rate, high school completion, years of tenure with the firm, age, median or above skill grade, and an indicator for speaking the local language (Kannada). Additionally, we check balance for several workplace outcomes: salary in the month before treatment announcement and productivity and task complexity in the announcement month (the first month of observation for these outcomes).

We fail to reject that the difference between treated and control workers for any of these outcome means at baseline is statistically significantly different from zero. Average attendance rates are about 90%, and average tenure with the firm is about 1.4 years. The average worker is about 27-28 years old. Over 60% of both samples are high school educated and speak Kannada.

The summary statistics and differences presented in Table 1 apply to the direct treatment comparison. Analogous balance checks for spillover comparisons were performed as well. We find no significant differences, and do not present them here for the sake of brevity.

[15] Risk and time preference modules were adapted from the Indonesian Family Life Survey.

Table 1: Summary Statistics

	(1)		(2)		(3)	
	Control		Treated		Difference	
P.A.C.E. Treatment	**Control Workers in Control Lines**		**Treated Workers in Treatment Lines**			
Number of workers	779		1,087			
	Mean	SD	Mean	SD	Mean Difference	p value
Attendance Rate (Jan-May 2013)	0.898	0.117	0.903	0.103	-0.005	0.380
High School	0.602	0.489	0.604	0.489	-0.003	0.901
Years of Tenure	1.432	2.709	1.353	2.119	0.079	0.500
Age	27.712	14.087	27.420	11.638	0.292	0.637
1(Speaks Kannada)	0.657	1.560	0.671	1.156	-0.014	0.834
High Skill Grade	0.616	0.843	0.642	0.688	-0.026	0.473
log(Salary) (May 2013)	8.746	0.188	8.737	0.156	0.009	0.258
Efficiency (Announcement Month)	0.586	0.587	0.556	0.426	0.030	0.268
SAM (Announcement Month)	0.618	0.726	0.615	0.535	0.003	0.928
Spillover Treatment	**Control Workers in Control Lines**		**Control Workers in Treatment Lines**			
Number of workers	779		837			

Notes: Tests of differences calculated using errors clustered at the line level according to the experimental design.

4 Empirical Strategy

4.1 Overview

The empirical analysis proceeds in several steps, beginning with testing the impact of the program on retention. This is important as a first step because impacts on retention would necessitate a weighting procedure to account for the differential attrition across treatment and control groups. Following this, we test for differences in workplace outcomes, then for differences in survey measures of self-reported personal and professional outcomes, and finally estimate treatment spillovers.

4.2 Retention, Working, and Cumulative Person Days

We estimate the following regression specification to test whether P.A.C.E. treatment impacts retention:

$$\begin{aligned} R_{wdmy} = \alpha_0 + \zeta_1 1[T_w] * 1[\text{Treatment Announced}]_{my} + \zeta_2 1[T_w] * 1[\text{During Treatment}]_{my} + \\ \zeta_3 1[T_w] * 1[\text{After Treatment}]_{my} + \psi_{uym} + \eta_w + \varepsilon_{wdmy} \end{aligned} \tag{1}$$

where the outcome is an indicator variable that takes the value 1 if worker w was retained on day d in month m and year y and 0 otherwise, $1[T_w]$ is a dummy variable that takes the value 1 if the worker is a trained worker on a treatment line and 0 if she is a control worker on a control line, and it is interacted with dummies that take the value 1 for the month that the assignment to treatment was announced, the months during the treatment and the months post-treatment, respectively, thus allowing comparison relative to the pre-announcement period. Each regression includes unit x year x month fixed effects ψ_{uym} (which absorb the main effects of the time dummies) and worker fixed effects η_w (which absorb the main effect of the treatment indicator).

We estimate equation 1 separately for retention dummy variables constructed using both daily attendance data and monthly payroll data. The difference between the two is that with the daily data we can see whether the worker stopped coming to work within the month, even before they are removed from the payroll. Standard errors are clustered at the production line level - while we did a two level randomized treatment assignment with the lower level of treatment at the worker level, we report line level clustering to be as conservative as possible. This is particularly important since we designed the experiment to measure spillover effects, and in fact find some evidence to this effect.

To estimate the impact of treatment on the additional number of days the firm receives from the worker, we consider two outcomes: the first is a binary working variable that is 1 if the worker was retained *and* is present in the the factory on a given day and 0 otherwise. It is thus a combination of retention and attendance. The second is the number of cumulative person days as measured by the cumulative running sum of the first variable. Both are defined at the daily level for each worker. They are estimated as in Equation 1 using these variables instead of retention on the left-hand side. These variables can once again be calculated from two sources of raw data: attendance and production rosters.

4.3 Dealing with Potential Bias from Selective Attrition

When examining conditionally observed outcomes such as productivity (which are only observed if the worker is still at the firm and working that day), there is a potential for selective attrition or observation based on treatment, which could generate bias in the impact estimates. To test and account for this potential bias, we follow several approaches, outlined below.

1. *Testing directly for treatment-induced changes in the relative size of treatment v. control groups:* We test directly for differential retention by estimating the regression specification in Equation 1 shown above. We present the results in Section 5.1. The results indicate there was no differential retention at the end point of the program period (July 2014) as well as any point afterward.

2. *Balance tests by baseline characteristics at different points during and post-program completion:* To test whether the retention across treatment and control is correlated with baseline characteristics, we present the results of balance tests by treatment and control one month after treatment (July 2014) as well as during the last month of data collection (February 2015). Results are presented in Table A9; the analysis shown here demonstrates that all baseline characteristics are balanced on means at both points in time. Tests conducted for other points in time are also balanced and omitted here for brevity. In addition, there is no heterogeneity in retention impacts across distributions of baseline characteristics at treatment announcement, program completion, and data collection endline, as shown in Figures A1-A6 (which provide a more stringent test than balance checks based on means).

3. *Dynamic weighting of conditionally observed outcomes:* As mentioned above, we do not find any differential retention at the end point of the program period, nor do we find any evidence of heterogeneity in retention across treatment and control groups for any baseline characteristics.

Despite this, in order to confidently recover population average treatment effects on conditionally observed outcomes throughout the observation period, we weight treatment and control groups by the probability of being observed at any intermediate point in the data. For example, if there exists differential attrition across treatment and control at 6 months into program implementation, even if this difference later equalizes, to ensure that we recover the population average treatment effect on any conditionally observed outcome (e.g., productivity or salary) at all subsequent points of observation, we can weight all observations prior to that time by the probability of being able to measure the outcome at each point in time. Accordingly, we adapt the approach proposed in Wooldridge (2010) to accommodate any potential heterogeneous impacts of treatment by baseline characteristics of the workers and any differential dynamics in the onset or decay of treatment effects across time, in the following manner:

(a) Estimate a probit specification for the probability of being observed, which is a dummy variable that takes the value 1 if the worker is in the sample on any given month and 0 otherwise (i.e., the retained dummy if studying impacts from the attendance or salary data and the working dummy if studying impacts from the production data), on the treatment indicator interacted with month by year fixed effects and baseline characteristics (attendance, education, tenure, age, skill grade, productivity and task complexity).[16]

(b) We then estimate equation 1 using the conditionally observed outcome variables on the left-hand side and the inverse of the predicted probabilities from the first step as probability weights. Note that because in the intermediate data (after the announcement but before the endline) the control group is less likely to be working (as shown in the results), this amounts to overweighting a subset of control observations at most points along the timeline.

In practice, once worker fixed effects are included in all regressions, the weighting procedure has negligible effect on the results. We explored robustness to different weights, as well as the absence of weights altogether, but do not present these results for the sake of brevity as they are generally quite similar.

4. *Production line-level estimates and impacts on retained workers only:* Finally, we present results for productivity and task complexity at the line level that includes all workers on the production lines, rather than at the individual level. Line level results are presented in Table A4 and discussed in detail in Section 6.5, and are quite consistent with individual-level results. (Note that we would expect smaller effects at the production line level, given that only a fraction of workers on each line were treated.) Additionally, estimates of productivity impacts for the subset of workers still retained by the end of the observation period are also reported in Table 3 and discussed in section 5.2 below. The pattern of results is the same for this subset of retained workers confirming that treatment impacts on productivity cannot be driven by changes in composition of the sample over time.

[16]Since workers salaries are homogenous within skill grade level, grade proxies for skill level as well as salary.

4.4 Productivity and Task Complexity

We estimate treatment impacts on two outcomes from the productivity data: efficiency and SAM. As discussed above, SAM measures task complexity, and efficiency is actual pieces produced divided by target pieces (calculated from SAM). All of these variables are only measured if a worker is retained by the factory, and present in the factory that day. Accordingly, these conditionally observed outcomes are weighted in the analysis as discussed above. The weights are obtained as discussed in section 3 using the working status dummy as the outcome.

In the SAM regressions, we follow the above specification exactly. However, in the efficiency regression, we replace the worker fixed effects with worker by garment style fixed effects. These are to account for any treatment impacts on the task complexity as identified in the SAM regression.

We also include as additional controls days that the style has been running on the production line and total order size to account for learning dynamics at the line level that might impact worker productivity across the life of the order.

4.5 Salary, Career Advancement, and Career Expectations

To study the impact of the program on career advancement, we measure impacts on gross salary and several work related survey outcomes. For salary, we first estimate the retention probability weights as detailed in section 3, and then estimate equation 1 using those inverse probability weights, with the log of gross salary as the outcome.[17]

We use five variables from the cross-sectional survey data to cover self-reported performance, subjective expectations of promotion, self-assessment, and initiative in requesting skill development. The subjective expectations of promotion were measured by a binary variable for whether the worker expects to be promoted in the next six months. The request for skill development was measured by asking workers whether they have undergone technical skill development training in the last six months. Self-reported performance was measured by asking whether workers have received production awards or incentives in the last 6 months. Finally, we measured two kinds of self-assessment. Both asked the worker to imagine a ladder with six steps representing the worst to best workers on their production line (6 being the best). The first self-assessment asked workers where they would place themselves relative to all the workers on their line, and the second where they would place themselves relative to other workers of their technical skill grade. Since the variation in the survey variables is only cross-sectional, we regress these outcomes on a binary variable for treatment or control, and include factory fixed effects, as well as control for age, tenure with the firm, and education of the worker. In survey outcome regressions, we employ weights obtained from the retention probit using attendance data matched to the date of survey.

4.6 Attendance, Unauthorized Leave, and Tardiness

We also analyze attendance outcomes, once again weighting these data by the inverse retention probabilities estimated from the probit specification discussed above. We focus the analysis on three out-

[17] Note that the administrative salary data is at the monthly level for each worker rather than the daily-level.

come variables: whether the worker is present at work, whether the worker is absent without leave (unauthorized) if absent, and whether the worker was tardy in coming to work.

4.7 Other Survey Outcomes

Finally, we consider the impact of the program on survey outcomes that might plausibly reflect the skills taught by P.A.C.E. For instance, since the program targets the stock of non-cognitive skills such as the ability to acquire and use information more effectively, we consider outcome variables regarding whether workers avail themselves of government and firm welfare programs like pension schemes and subsidized health-care. Similarly, since the program aims to make workers more forward-looking, we test whether there is an increase in workers' savings, especially for important future considerations like education (their own or their children's), and risk and time preferences. Furthermore, we test whether the program impacted personality characteristics (conscientiousness, locus of control, perseverance, extraversion and self-sufficiency) and mental health (self-esteem, hope/optimism, and mental distress.). As mentioned previously, the survey measures are cross-sectional. The regression specification is thus the same as for the survey outcomes in the previous section: we regress the outcome on the binary treatment variable and include factory unit fixed effects and retention weights from the attendance data matched by survey date.

4.8 Figures

We create figures illustrating the month-by-month treatment impacts by re-estimating all the outcome regressions with the treatment binary interacted with monthly dummies from June 2013 onwards (rather than the announcement, during, and after dummies presented in equation 1 above). All regression analogs are reported in tables in the Appendix, with figures presented and discussed in section 5. Dummies for months prior to June 2013 are excluded to make treatment effects relative to the pre-announcement period in all figures, except for those depicting monthly treatment impacts on productivity outcomes for which the announcement month (June 2013) is the first month of observation and the excluded base month.

4.9 Spillover Effects/Production Complementarity Effects

To estimate the effects on untrained workers who interact with trained workers, we re-run all of the specifications mentioned above, replacing the binary treatment variable with the binary spillover treatment variable. This variable compares untrained workers in treatment lines (workers who enrolled in the lottery but did not receive the program and who work in production lines with workers who received the training) with control workers in control lines (workers who enrolled in the lottery but did not receive the program and who work in production lines without any trained workers). Thus, it takes the value 1 if the individual is an untrained worker in a treated line, and 0 if the worker is a control worker in a control line (and missing for treated workers).

5 Results

5.1 Retention and Daily Working Status

Figure 2: Monthly Retention

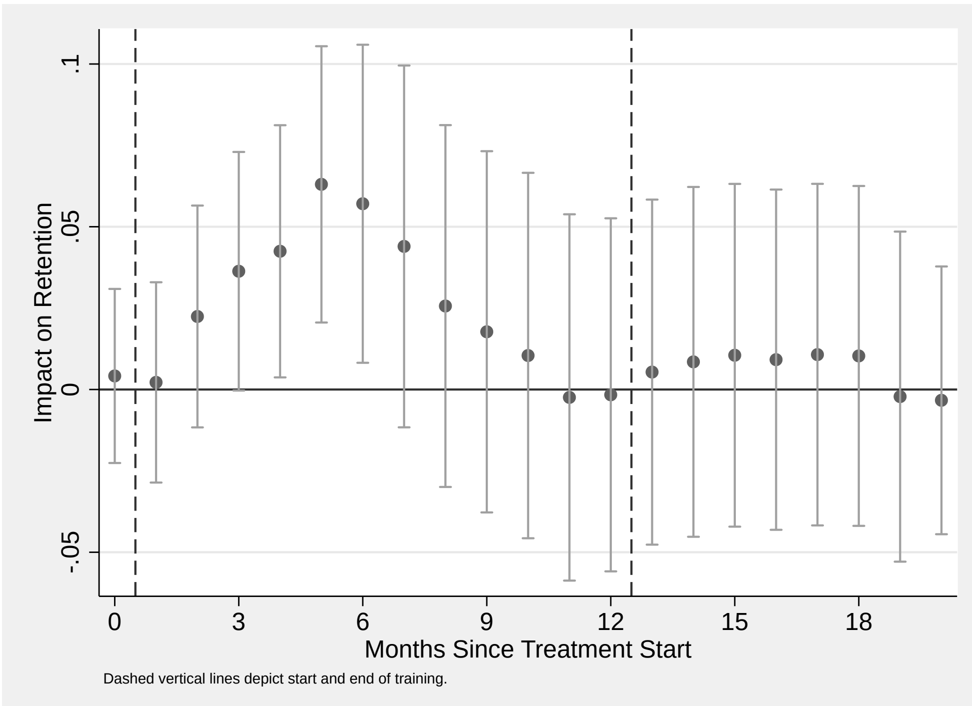


Figure 2 depicts impacts of P.A.C.E treatment on retention. Figure 2 plots coefficients of monthly impacts from the preferred regression specification. The corresponding full results are reported in Table A2 in the Appendix. Figures using payroll roster data instead of attendance data look nearly identical. Accordingly, these are not presented, but are also available upon request. Table 2, however, does present analogous regression results from all of these alternative samples. Figure A7 in the Appendix depicts raw retention data from the attendance roster across P.A.C.E treatment and control groups over the full observation period.

We begin by measuring the impacts of P.A.C.E. on retention and the probability that a worker is on the job.[18] Figure 2 plots regression coefficients of treatment effects estimated month by month using attendance roster data. This figure shows that there is a statistically significant impact of treatment on retention early in the program period, which dissipates by the end of the program (the program training window is denoted by dashed vertical lines). Column 1 of Table 2 presents analogous regression coefficients pooling months after program assignment into three periods: announcement, during training, and after training. The results indicate that on average the treatment impacts on retention were small and insignificant throughout the entire observation period. Using the payroll data yields nearly identical figures and so this additional figure is omitted for brevity. We do, however, report estimates using this alternate data in column 2 of Table 2.

[18]Since all the variables discussed in this section are not conditional on retention (i.e., not missing if the worker has left the firm), no re-weighting is required.

Figure 3: Monthly Working

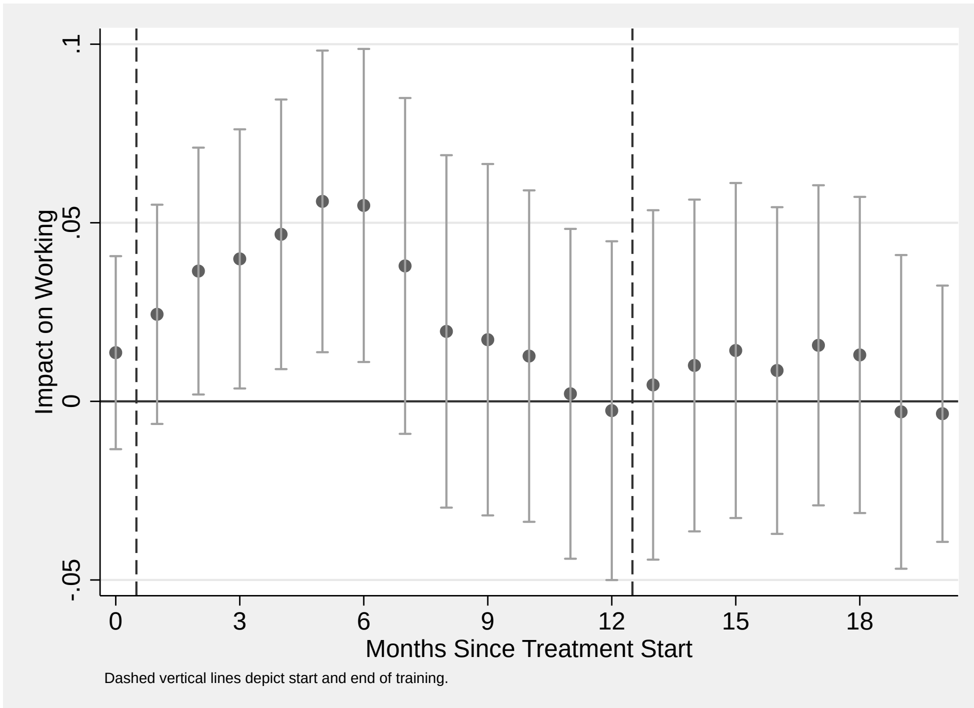


Figure 3 depicts impacts of P.A.C.E treatment on working (retained and present) in the factory from the attendance roster data. Figure 3 plots coefficients of monthly impacts from the preferred regression specification. The corresponding full results are reported in Table A2 in the Appendix. Figure A8 depicts raw presence data from the attendance roster across P.A.C.E treatment and control groups over the full observation period.

The second outcome of interest is the probability that a worker is retained *and* present at work on a given day. This variable, which we refer to as "working" status, is therefore equal to 0 on a given day if the worker has permanently left the factory, or she is still working for the firm but is not present on a given day, and is 1 otherwise. Figure 3 plots regression coefficients of month-by-month treatment effects for the attendance roster data. Figure 3 once again shows that treatment impacts are statistically significant for some of the treatment period but not afterward (the program period is once again denoted by dashed vertical lines).

Columns 3 and 4 of Table 2 present analogous regression coefficients for pooled post program assignment months. Treatment impacts are large and significant during and after the program when using production roster data, but attenuated and imprecise when using attendance data. This difference is likely due to measurement error from two sources: 1) attendance data is more prone to measurement error when biometric scanning equipment is malfunctioning or workers forget to scan in at the start of the day;[19] and 2) attendance data records partial-work days as absences where as production data will count workers as present if they record production in that day.[20] In the production data, a worker was

[19]This is particularly salient for treatment workers during the program months as the training spans the usual time in the morning when workers would scan in for the day.

[20]The means of the control group across the two sources are different due to the fact that the production data is only

Table 2: Impacts of P.A.C.E. Treatment on Retention, Working, and Person Days

	(1)	(2)	(3)	(4)	(5)	(6)
	Retained		Working		Cumulative Person Days	
	1(Worker Still on *Attendance* Roster)		1(Worker Retained and Present in Factory Today)		Sum of Days Working for Each Worker to Date	
	Attendance Roster	***Payroll Roster***	***Attendance Roster***	***Production Data***	***Attendance Roster***	***Production Data***
After X P.A.C.E. Treatment	0.00620	0.00865	0.00743	0.0761**	9.250	16.20**
	(0.0256)	(0.0274)	(0.0221)	(0.0371)	(8.683)	(7.141)
During X P.A.C.E. Treatment	0.0264	0.0256	0.0285	0.0870***	5.360	6.833***
	(0.0215)	(0.0220)	(0.0193)	(0.0318)	(3.258)	(2.601)
Announced X P.A.C.E.. Treatment	0.00416	0.00476	0.0136		0.501	
	(0.0136)	(0.0153)	(0.0138)		(1.271)	
Fixed Effects			Unit X Month X Year, Worker			
Observations	1,433,981	43,141	1,270,871	778,916	1,270,871	778,916
Control Mean of Dependent Variable	0.63	0.66	0.52	0.37	213.71	103.22

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the treatment line level. Retained dummy, Working dummy, and Cumulative Person Days are all defined for every worker date observation in the data and therfore the regressions do not require any weighting.

8.7 percentage points more likely to be working during the program (a 23.5% increase relative to the control mean) and 7.6 percentage points more likely to be working after the program (a 20% increase relative to the control mean).

The final measure we study regarding retention and working status is the cumulative number of working days that accrue to the firm. This is the running sum of the working status variable just discussed. Figure 4 shows that the treatment impact on cumulative person days (calculated from production data) is positive and statistically significant by about 3 months into the program period. The impacts continue to grow quickly through month 8 of the training period, after which the growth slows somewhat but remains positive through the remainder of the observation period. Columns 5 and 6 of Table 2 present the impacts on cumulative person days during and after the program, using attendance and production data, respectively. The treatment increases the cumulative person days per treated worker by 6.8 days during treatment and 16.2 days after treatment when the production data is used, which is about 6.6% and 16% of the mean cumulative number of days of the control group respectively.

5.2 Productivity and Task Complexity

If P.A.C.E. impacted the stock of soft skills (e.g., time management, communication, extraversion), then it should follow that marginal productivity rises, both through direct channels, to the extent that soft skills are used in production, and indirect channels, if workers were more likely to ask for and receive additional training in hard skills. To test this hypothesis, we consider two outcomes: 1) productivity as reflected in the industry standard measure of efficiency (pieces produced divided by target pieces);

available starting June 2013 (the month of treatment announcement), so has five months less of data relative to the attendance roster.

Figure 4: Monthly Person Days

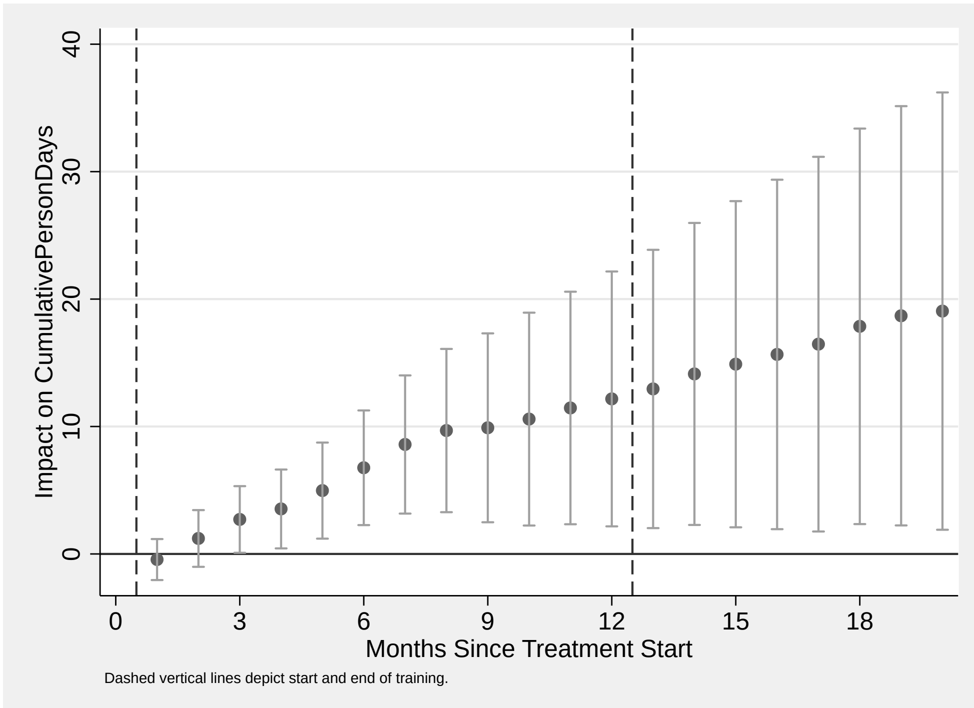


Figure 4 depicts impacts of P.A.C.E treatment on cumulative person days in the factory from the start of the production data (June 1, 2013) to each date. Figure 4 plots coefficients of monthly impacts from the preferred regression specification on the production data. The corresponding full results are reported in Table A2 in the Appendix. Figure A9 depicts raw person days data from the production data across P.A.C.E treatment and control groups over the full observation period.

and 2) the complexity of the task to which workers are assigned, as measured by SAM (number of minutes in which a task is expected to be completed – a higher SAM thus denotes a more complex task).

Figures 5A and 5B plot regression coefficients of impacts of treatment on efficiency, estimated month by month. Figure 5A presents this for all workers in the sample, and Figure 5B for only those workers who were retained at the end of the data collection period (February 2015). The figures indicate that treatment increases efficiency throughout the training and post-program period, with coefficients becoming significant towards the last third of the program period and after. Figures 6A and 6B plot analogous regression coefficients of monthly treatment impacts on the complexity of the operation the worker is performing as measured by SAM. These figures illustrate that both during and after the program, there is evidence that treated workers are assigned to more complex tasks (tasks with higher SAM).

These patterns are confirmed in Table 3, which reports the results of analogous regressions in which impacts are grouped into during and after P.A.C.E. program implementation. Treated workers are more efficient after the program (relative to the month of treatment assignment announcement) by nearly 11 percentage points, about 20% relative the control group mean. Consistent with the evidence presented above, we see that the impacts on productivity are stronger after program completion. For

Figure 5A: Monthly Efficiency (All Workers)

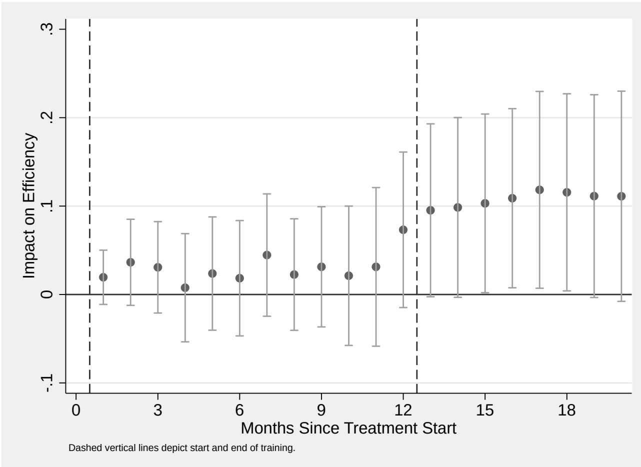
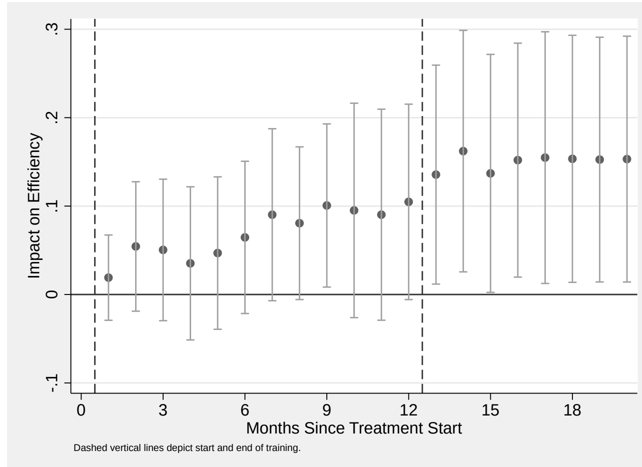


Figure 5B: Monthly Efficiency (Retained Only)



Figures 5A and 5B depict impacts of P.A.C.E treatment on productivity in the factory. Figure 5A depicts coefficients of monthly impacts on efficiency (actual pieces produced / target pieces) from the preferred regression specification (including worker by item (style) fixed effects and controls for the number of days the worker has been producing that style on that line and the total order quantity) for the full sample of workers, with observations weighted to account for any differential composition across treatment and control due to attrition. Figure 5B presents the analogous figure for the subsample of workers who are still retained in the factory by the end of observation (February 2015). The corresponding full results are reported in Table A3 in the Appendix.

Figure 6A: Monthly SAM (All Workers)

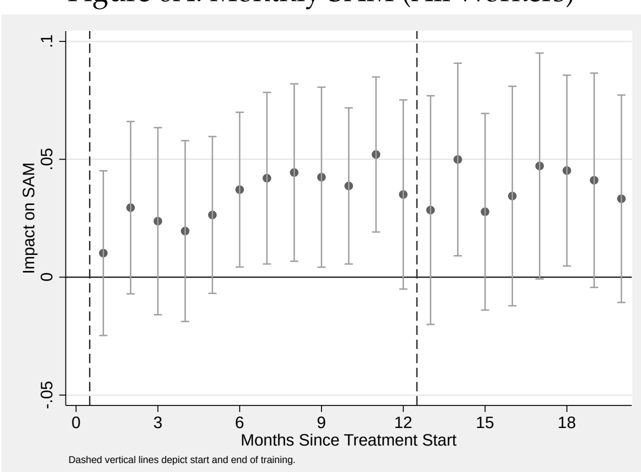
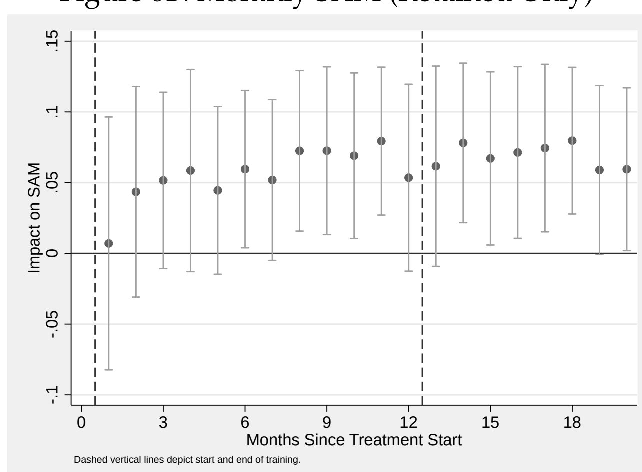


Figure 6B: Monthly SAM (Retained Only)



Figures 6A and 6B depict impacts of P.A.C.E treatment on operation complexity (SAM, or standard allowable minute per operation-piece). Figure 6A depicts coefficients of monthly impacts from the preferred regression specification for all workers. Figure 6B depicts monthly impacts for the subsample of retained workers only. The corresponding full results are reported in Table A3 in the Appendix. Figure A11 depicts raw SAM from the production data across P.A.C.E treatment and control groups over the full observation period (June 1, 2013 onwards in the production data).

the sub-sample of workers who were retained until the end of the data collection period, the magni-

Table 3: Impacts of P.A.C.E. Treatment on Productivity

	(1)	(2)	(3)	(4)
	Efficiency Produced/Target	SAM (Operation Complexity) Standard Allowable Minute	Efficiency Produced/Target	SAM (Operation Complexity) Standard Allowable Minute
			Retained Workers Only (still in factory in Feb 2015)	
After X P.A.C.E. Treatment	0.108**	0.0384**	0.150**	0.0798***
	(0.0510)	(0.0180)	(0.0654)	(0.0255)
During X P.A.C.E. Treatment	0.0300	0.0334**	0.0693*	0.0642***
	(0.0274)	(0.0147)	(0.0390)	(0.0208)
Additional Controls	Days on Same Line-Garment, Total Order Size	None	Days on Same Line-Garment, Total Order Size	None
Fixed Effects	Unit X Month X Year, Worker X Garment	Unit X Month X Year, Worker	Unit X Month X Year, Worker X Garment	Unit X Month X Year, Worker
Weights	Inverse Predicted Probability from Probit of Working on Treatments X Mo-Yr X Baseline Characteristics		None	
Observations	290,763	290,763	130,187	130,187
Control Mean of Dependent Variable	0.542	0.565	0.527	0.588

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the treatment line level. Observations in columns 1 and 2 are weighted in regressions by the inverse of the predicted probability of working (i.e., not yet attrited and present in the factory with non-missing data) in the sample that day from a probit regression of the working dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. Sample in columns 3 and 4 is restricted to only workers still retained in the factory by the end of observation. All samples are trimmed in these regressions to omit days in which the worker is observed for only a half a production day or less or days in which the worker is observed for more than 2 overtime hours as these are anomalous observations with imprecise production measures. These outliers make up only around 5% of the work-day observations.

tude of the treatment effect is similar, about 15 percentage points higher efficiency after the treatment. Figure A10 presents these coefficients for the whole sample and the subsample of retained workers only together as well as their confidence intervals to test for statistically differences in every month of data collection. We cannot reject that the coefficients are the same in any month. The fact that these results are similar across panels further supports the notion that any changing composition of the sample can be driving the productivity impacts.

Additionally, we see fairly consistent impacts on task complexity (SAM) throughout the program, and they are sustained and remain statistically significant after the program period. That is, treated workers are assigned to more complex tasks both during and after treatment (tasks to which they are assigned are expected to take about 2.3 seconds (0.038 minutes) more, roughly 7% of the control group mean). Thus, not only are workers in the treatment group assigned to more complex tasks during and after the program, they are more productive even at these harder tasks once treatment ends. The non-cognitive skills that the program covers (like time management, goal setting, and team work) enhance worker productivity and the ability to perform complex tasks.

The time pattern of impacts on productivity – insignificant increases during much of the program period followed by large, significant increases towards the end of training and afterward – is striking and deserves additional consideration. The observed pattern could be rationalized in several ways. First, the increase in task complexity discussed above (which happened early on in the program period) may not captured fully by adjusting the target quantity. More complex tasks may take longer to master, creating a drag on efficiency particularly just after task switching occurs. Second, the "incubation period" for productivity impacts in the context of this program, through both direct and indirect channels mentioned above, is likely long. Learning soft skills to the point that they can be applied in the workplace may take time. Third, Sets of soft skills may be complementary, so that incremental learnings in a given module have a greater impact later in the program. This is consistent with the structure of the program, which conducted review sessions before graduation to reiterate earlier modules and discuss how to combine the new skills together and apply them in both professional and personal situations. Finally, from anecdotal observation, women took several months to become true participants in the group sessions; at the beginning of the program the level of participation, fitting with the cultural context in which these women live, was quite low.

5.3 Career Advancement

In addition to worker presence and productivity, we study career advancement within the firm. To estimate the impacts of treatment on career advancement, we consider both whether the worker was given a raise using monthly payroll data as well as worker-reported measures of expectations of promotion; whether they recently asked for (and received) skill development training; earned production incentives; and finally, how they assess their own ability relative to all workers on their production line, and relative to workers of the same technical skill grade as them. Except for the salary data which is at the monthly level for each worker, the self-reported measures are from the worker-level survey conducted in the month of program completion and vary only cross-sectionally.

Figure 7: Monthly Salary

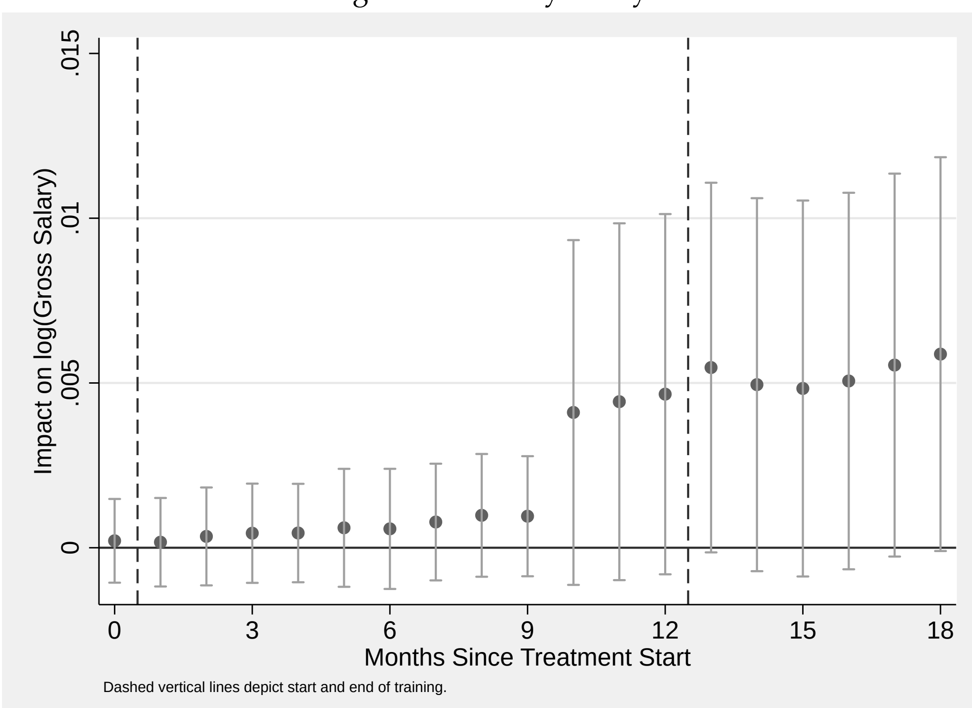


Figure 7 depicts coefficients of monthly impacts of P.A.C.E treatment on log(gross salary) from the preferred regression specification. The corresponding full results are reported in Table A3 in the Appendix.

Figure 7 plots regression coefficients of monthly treatment impacts on log gross salary. We see in

Figure 7 that PACE workers are paid negligibly more (roughly half a percent), with the gap showing up towards the end of the program period and trending up modestly thereafter. Column 1 of Table 4 presents the results of the estimation comparing treatment workers to control workers during the treatment assignment announcement month, and during and after the treatment (relative to before the treatment assignment announcement month). Treatment workers receive on average less than half a percent more wages in the period after the program completion, which translates to roughly 30 INR or less than .5 USD a month. Thus, despite being assigned to more complex tasks and being more productive, treated workers are not paid meaningfully higher wages.

Columns 2-6 of Table 4 presents the results from analysis of related survey outcomes. Treatment workers are about 8.7 percentage points more likely to report that they expect a promotion within the next six months (roughly 15% of the control group mean), and are nearly 16 percentage points more likely to request skill development training (63% of the control group mean). They are not significantly more likely to report having received a production incentive or award, but rate themselves higher relative to peer co-workers. Specifically, when asked to rank themselves relative to workers the same technical skill grade, they are significantly more likely to rate themselves at a higher level (as shown in column 5).

5.4 Attendance

Related outcomes of interest are attendance (a binary variable that is 1 if the worker is at work today and 0 if not), unauthorized leave (a binary variable that is 1 if the worker is not at work today and did not inform the employer and 0 if she is either at work or absent and took prior formal leave from the employer), and tardiness (a binary variable that is 1 if the worker was late relative to the modal arrival time of co-workers on the line that day and 0 if not). Table A5 presents the impacts of treatment on these outcomes. There are no precisely measured impacts on any of the outcomes if the grouping is done by these milestones rather than a month by month comparison. Table A6 presents the regression results of the month by month estimation. The results indicate that treatment workers are more likely to attend work in the first two months of the program, and absences are more likely to be authorized during the same months. Worker tardiness does not appear to be impacted during or after treatment.

6 Mechanisms

Our interpretation of the productivity and task complexity results is that skills like time and stress management; communication; problem solving and decision-making; and effective teamwork are "soft" inputs into production. Reinforcing these skills through the P.A.C.E. program should thus directly affect workplace outcomes. Across the categories of results presented below, impacts are consistent with a direct treatment effect on the stock of soft skills. In particular, the narrative that emerges is one that is consistent with the P.A.C.E. program increasing the stock of soft skills. This is indicated in part by the fact that treated women are more likely to proactively increase their stock of hard skills by requesting technical training, are more extraverted, more likely to seek out and avail themselves of government and employer benefits to which they are entitled, and more likely to exhibit forward looking behav-

Table 4: Impacts of P.A.C.E. Treatment on Salary and Workplace Related Survey Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(Gross Salary)	Expect Promotion Next 6 Mos	Skill Development Training	Production Award or Incentive	Skill Peer Self-Assessment	Co-Worker Self-Assessment
	Salary Data			***Survey Data***		
After X P.A.C.E. Treatment	0.00492*					
	(0.00270)					
During X P.A.C.E. Treatment	0.00137					
	(0.000906)					
Announced X P.A.C.E. Treatment	0.000221					
	(0.000647)					
P.A.C.E. Treatment		0.0871**	0.158***	0.0293	0.122*	0.0645
		(0.0414)	(0.0467)	(0.0185)	(0.0648)	(0.0667)
Fixed Effects	Unit X Month X Year, Worker			Unit, Education, Age, Tenure		
Weights	Inverse Predicted Probability from Probit of Retention on Treatments X Mo-Yr X Baseline Characteristics					
Observations	28,692	621	621	621	621	621
Control Mean of Dependent Variable	8.909	0.563	0.249	0.032	5.337	5.298

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the line level. Observations are weighted in regressions by the inverse of the predicted probability of being retained (i.e., not yet attrited with non-missing data) in the sample that day from a probit regression of the retained dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. Controls in regressions for survey outcomes include demographic baseline variables from Table 1 (i.e., dummies for education levels, dummies for deciles of the age distribution, and dummies for tenure in integer years).

ior via savings and aspirations for their children's future. Finally, these women share learnings with their untrained co-workers, and these spillovers appear to contribute to the productivity of these co-workers.

Below, we support this interpretation using evidence from a survey of treatment and control workers; from assessments of the treatment group's knowledge before and after the completion of the program's core modules; and from the degree of treatment spillovers. We also present several alternative interpretations and discuss the plausibility of each in turn.

6.1 Survey Results

The first piece of evidence supporting the interpretation that the stock of soft skills changed comes from a survey we administered to treatment and control workers in the month after program completion. Table 5 evaluates the impact of P.A.C.E. treatment on financial behaviors and attitudes (Panel A); availing of firm and government programs (Panel B); personality (Panel C); and mental wellbeing and aspirations (Panel D).

We discuss results within each category in turn to lay out our reasoning. The first category is meant to evaluate whether P.A.C.E. treatment changes women's financial behaviors and attitudes. This change would be consistent with a shift in forward-looking behavior, an important dimension of soft skills. The results from Panel A indicate that there is a positive impact on saving for own and children's education, and the impacts are quite large relative to the control group mean (about 30% of the control group mean). Savings for other purposes show no significant impacts. We construct survey-based measures of risk-aversion and patience with higher scores corresponding to higher levels of risk aversion and patience. The estimates suggest that treatment increases risk aversion as well as participation in insurance or informal risk-sharing mechanisms (about 10% from the control group mean).

The second category, availing oneself of government and employer-based entitlement programs, is meant to evaluate changes in the effectiveness of information acquisition, another important soft skill. The results in Panel B show that treated workers are substantially more likely to seek out welfare programs. Impacts on binary indicators for enrollment in government pension and government subsidized healthcare indicate that treated workers are more likely to avail themselves of these programs. The magnitude of these impacts are quite large relative to control group means, which are around 0 for both outcomes. Impacts on other government subsidies and firm entitlements are negligible.

The third category, personality, is meant to assess differences in key traits that are associated with personality traits, namely conscientiousness, locus of control, perseverance, extraversion, and self-sufficiency. In general, the impact estimates (shown in Panel C) are imprecisely estimated, but P.A.C.E. treatment does have a large positive and statistically significant impact on extraversion. This result on extraversion is consistent with the results above related to seeking out information and resources, as well as results on self-reported comparisons to co-workers, which show that P.A.C.E. training increased self-regard with respect to workplace performance relative to peers.

The final category of the survey, mental health and aspirations, is meant to assess impacts on psychological well-being and the extent to which future aspirations are affected by treatment. The results

Table 5: Impacts of P.A.C.E. Treatment on Survey Outcomes

	(1)	(2)	(3)	(4)	(5)
Panel A: Financial Behaviors and Attitudes	Saving for Education	Saving for Other Reasons	Risk Preference Index	Time Preference Index	Insurance or Informal Risk-Sharing
P.A.C.E. Treatment	0.0804**	-0.0465	0.166*	-0.0984	0.0637*
	(0.0313)	(0.0334)	(0.0876)	(0.0935)	(0.0351)
Control Group Mean of Dependent Variable	0.265	0.272	-0.052	0.019	0.628
Panel B: Government and Firm Entitlements	Gov. Pension	Gov. Subsidized Healthcare	Other Gov. Subsidy	Firm Entitlements	Community Self Help Group
P.A.C.E. Treatment	0.0248*	0.0226**	0.0119	-0.0257	-0.0270
	(0.0141)	(0.00941)	(0.0310)	(0.0352)	(0.0303)
Control Group Mean of Dependent Variable	0.039	0.006	0.120	0.142	0.152
Panel C: Personality	Conscientiousness	Locus of Control	Perserverance	Extraversion	Self-Sufficiency
P.A.C.E. Treatment	0.0210	0.0307	-0.123	0.164**	0.0445
	(0.0732)	(0.0770)	(0.0774)	(0.0702)	(0.0877)
Control Group Mean of Dependent Variable	-0.047	-0.040	0.020	-0.071	-0.063
Panel D: Mental Health and Aspirations	Self-Esteem	Hope/Optimism	Moderate Distress	Child's Expected Age at Marriage	Child Educated Beyond College
P.A.C.E. Treatment	-0.172	-0.0621	-0.0422	0.0456	0.0885***
	(0.106)	(0.0819)	(0.0389)	(0.165)	(0.0280)
Control Group Mean of Dependent Variable	0.048	0.015	0.094	23.427	0.117
Fixed Effects	Unit, Education, Age, Tenure				
Weighted	Inverse Predicted Probability from Probit of Retention on Treatments X Baseline Characteristics				
Observations	621	621	621	621	621

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the treatment line level. Obersvations are weighted in regressions by the inverse of the predicted probability of being retained (i.e., not yet attrited with non-missing data) in the sample that day from a probit regression in the attendance roster of the retained dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. Controls include demograhpic baseline variables from Table 1 (i.e., dummies for education levels, dummies for deciles of age distribution, and dummies for tenure in integer years).

reported in Panel D show that, in general, outcomes associated with psychological well-being (self-esteem, optimism, and mental distress) are unaffected by P.A.C.E. treatment, but aspirations for children's education rise dramatically in relation to the control group mean. This is consistent with the result on saving for education presented in Panel A.

6.2 Pre- and Post-Module Assessments

Figure 8A: Avg Pre Score

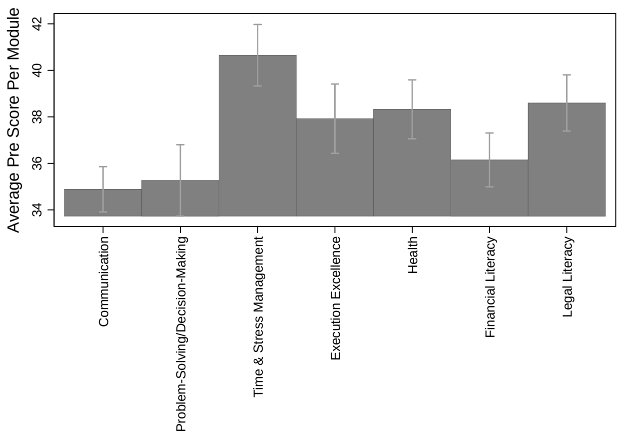
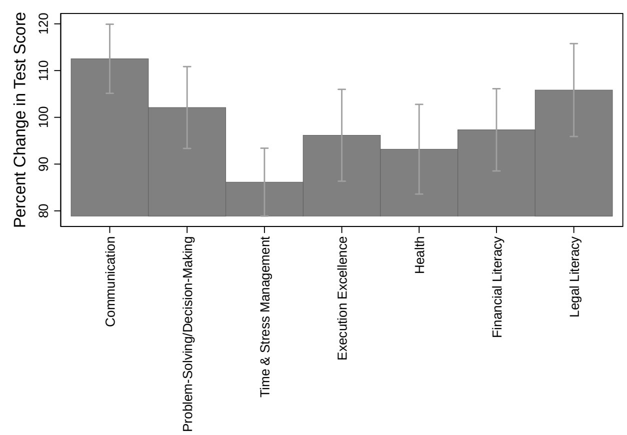


Figure 8B: Post - Pre % Change in Score



Figures 8A and 8B depict average pre-training test score (8A) and normalized (percent change, 8B) difference between post- and pre-training test scores administered for all core P.A.C.E. modules. Raw scores for each assessment are out of 100. These assessments were not given to control workers and accordingly cannot be analyzed in the preferred specification. Figure A12 in the appendix shows average session attendance rates by training module.

The second source of evidence on the direct impacts of P.A.C.E. on the stock of soft skills is pre- and post-module assessments built into the program. These assessments were designed to test the specific value added from each core program module. They were only administered to program participants, and thus we cannot compute a treatment vs. control difference, rather only a post vs. pre-module difference for treated workers.

Figure 8A shows the pre-module assessments for each core P.A.C.E module. Figure 8B shows the percent change between (identical) assessments taken pre- and post-module for each core P.A.C.E. module. Taken together, the results from both analyses show that P.A.C.E. participants had low baseline stocks of soft skills and improved their stocks of these skills dramatically through the training. The changes shown in Figure 8B are all in the neighborhood of 85-110 percent, with the largest changes (in percent terms) for Communication, Problem Solving/Decision-Making, Legal Literacy, and Execution Excellence. The largest raw difference is in the Time and Stress Management module.

These results support the notion that workers absorbed the skills taught in each of the core modules, and that the stock of skills increased. We should note some caveats in interpreting these changes. First, as described above, control workers were not given the assessments, so we are not able to estimate impacts by comparing across the treatment and control groups). Second, we are measuring skill retention directly after module completion; this does not necessarily reflect long-term skill retention. Nevertheless, these results are consistent with our hypothesis that P.A.C.E. acted on workplace outcomes by increasing the stock of soft skills.

6.3 Treatment Spillovers

Finally, we consider evidence on spillovers. Recall that the experiment was designed to capture spillovers within production lines through a two-stage randomization procedure, in which lines were first randomized to treatment or control, and then within treatment lines, workers who had enrolled in the P.A.C.E. lottery were randomized to treatment or to the spillover group. In this section we evaluate spillovers by comparing the outcomes of this latter group to control workers on control lines. The existence of spillovers would provide greater justification for employer investment in soft skills training. We evaluate these hypotheses in Table 6, which presents the spillover results for workplace outcomes of interest.[21]

Panel A presents the results for person days as well as productivity. There is a weakly statistically significant impact on the binary for working during the treatment announcement period, and a stronger result for cumulative person days during the treatment period - untrained workers who work with treated workers work for about 8 more days during program months relative to control workers. Productivity impacts are positive, about 70% as large as the direct treatment effects, but are not statistically significant. Panel B presents the results for career advancement variables. Similar to the effect on productivity, the spillover impacts on survey outcomes on requesting skill development training, receiving a production incentive or self-assessment relative to co-workers are not precisely measured, but again have coefficients of the same sign as the main treatment impacts. The worker self-assessment relative to co-workers is positive and statistically significant at the 10% level.

Table A8 presents the results for the non-workplace outcomes of interest for spillovers. On the whole, estimates for non-workplace outcomes do not show a strong pattern of spillover impacts. However, like directly trained workers, spillover workers are also more likely to be saving for their childrens' education and utilizing government subsidized healthcare.

In sum, then, for workplace outcomes we see large spillover impacts on cumulative person days accrued to the firm, and imprecisely estimated but positive effects on efficiency. We see some evidence for spillovers on outcomes outside the workplace, but the results are imprecise in general. Overall, the presence of spillovers suggests that knowledge transfer happened as a direct result of the program – i.e., that program participants imbibed soft skills, which they then communicated to co-workers on their production lines, and that transfer helped improve outcomes of non-participants, as well.

6.4 Alternative Mechanisms

Having presented evidence on the salience of direct skilling as a result of the P.A.C.E. program, we now discuss several alternative interpretations of the results and any supporting evidence of each.

[21]Note that probability weights, when necessary, are calculated exactly as they are in the treatment effect estimation, using spillover treatment indicators in place of direct P.A.C.E. training.

Table 6: Spillovers on Co-Workers (Attendance, Productivity, and Career Advancement)

	(1)	(2)	(3)	(4)	(5)
Panel A: Working and Production	Working		Cumulative Person Days		Efficiency
	Attendance	***Production***	***Attendance***	***Production***	
After X Spillover	-0.0155	0.0363	8.652	8.092	0.0714
	(0.0206)	(0.0438)	(9.332)	(7.884)	(0.0571)
During X Spillover	0.0252	0.0628	8.023**	4.751	0.00591
	(0.0209)	(0.0386)	(3.841)	(3.050)	(0.0319)
Announced X Spillover	0.0317*		2.151		
	(0.0172)		(1.372)		
Fixed Effects	Unit X Month X Year, Worker				Unit X Month X Year, Worker X Garment
Weights	None				Inverse Predicted Probability from Probit of Working on Treatments X Mo-Yr X Baseline Characteristics
Observations	1,102,880	673,407	562,478	673,407	241,322
Control Mean of Dependent Variable	0.519	0.382	0.390	107.437	0.548

Panel B: Career Advancement	Skill Development Training	Production Award or Incentive	Skill Peer Self-Assessment	Co-Worker Self-Assessment
Spillover	0.0254	0.0204	0.113	0.140*
	(0.0608)	(0.0243)	(0.0687)	(0.0769)
Fixed Effects	None			
Weights	Inverse Predicted Probability from Probit of Retention on Treatments X Mo-Yr X Baseline Characteristics			
Observations	527	527	527	527
Control Mean of Dependent Variable	0.244	0.031	5.287	5.267

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the treatment line level. All regressions are for sewing department workers only as spillover sample is not defined for non-sewing workers. Retained and working dummies and cumulative man days are defined for every worker date observation in the data and therfore regressions do not require any weighting. Observations in attendance and advancement regressions are weighted in regressions by the inverse of the predicted probability of being retained (i.e., not yet attrited with non-missing data) in the sample that day from a probit regression of the retained dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. Controls for survey outcome regressions in Panel B include demograhpic baseline variables from Table 1 (i.e., dummies for education levels, dummies for deciles of age distribution, and dummies for tenure in integer years).

First, we address the potential importance of reciprocity (an impulse to give back to the employer as a result of access to the program). While it is plausible that some part of the impacts observed is due to reciprocity, we deem it unlikely that the majority of impacts are due to this mechanism. This is for two reasons. First, we find spillovers in treatment for the number of days worked by workers who were signed up for the program and were on the same production line as treatment workers, but did not receive the program. These would be difficult to explain if reciprocity were the main driving force behind workplace impacts, since non-participants should not be driven by this motive. Second, productivity impacts accumulate slowly during the program period and persist strongly for at least 8 months after program completion, with the largest productivity impacts occurring during this post-training period. This does not fit well with a reciprocity motive as a primary mechanism, since we would expect the reciprocity motive to be strongest while the program is offered and to dissipate over time if pay does not rise commensurately with productivity as in this case. This indirect evidence is in line with recent, more direct tests of the role of reciprocity in workplace settings (DellaVigna et al., 2016).

Second, we evaluate the possibility that the results for productivity and task complexity were due to sheepskin effects, i.e., taking part in P.A.C.E. "certified" workers as high quality from the perspective of management, and this led to the improvements in workplace outcomes we observe. We reason that sheepskin effects are unlikely to explain the majority of the program's impacts given the slow onset of increased productivity over time, rather than an increase near the program's end. Additionally, once again spillover impacts are inconsistent with a sheepskin effect mechanism.

Third, it is possible that workers found the classes enjoyable and they improved workers' subjective wellbeing, which in turn made workers more productive. The results reported in Panel D in Table 5 show that levels of psychological distress are unaffected by treatment, which contradict changes in worker well-being and happiness being the mechanism for productivity impacts.[22]

Finally, we evaluate the idea that increased social capital drives the results on workplace impacts. The argument here is that it is possible that P.A.C.E. sessions improved the ability of workers to create social ties, which could generate higher productivity on their production lines if it increased the extent or intensity of social connectivity on the line. We argue that the context in which the study was run likely precludes this from being a primary mechanism of impact. First, language-based and cultural barriers are quite salient in the workplace in our context, likely limiting the extent of the importance of social connectivity in productivity. Nearly half the workers in the factories under study are migrants, many of whom do not speak Kannada, the indigenous language of Karnataka. Second, due to throughput constraints which dictated the number of workers from the same production line who could leave at the same time for a P.A.C.E. session, co-workers on the same line were placed in different sessions conducted on different days of the week. Again, this likely limited the increase in within-line social connectivity. These explanations do not preclude overall social ties from being impacted by the program; they simply lower the likelihood that this channel contributed significantly to impacts on workplace outcomes like productivity.

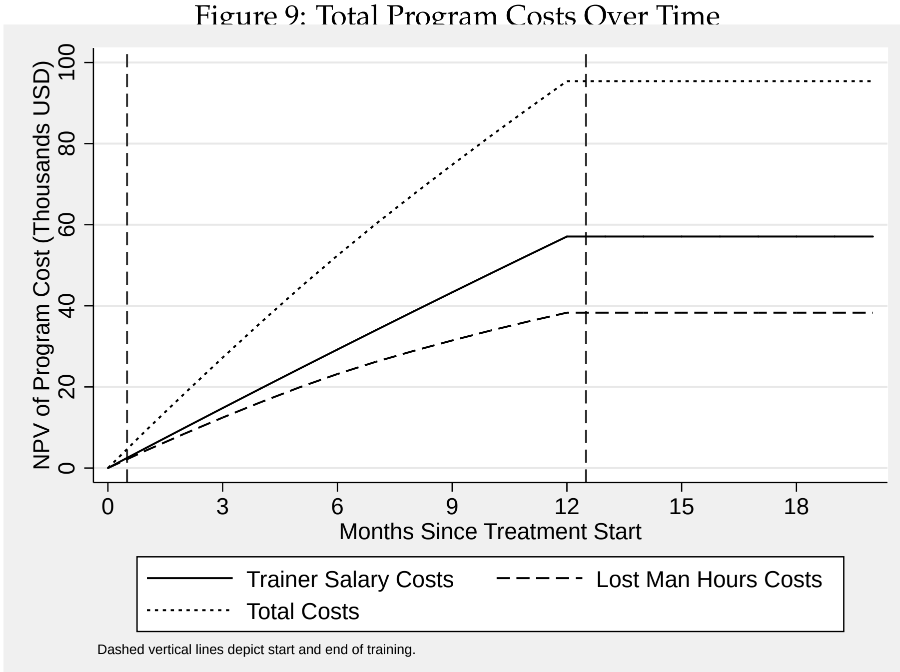
[22] Results are unchanged if severe mental distress is used as an outcome instead of moderate mental distress.

6.5 Line-Level Productivity and Task Complexity Results

As a further test of robustness of our main results, we present regression results using daily productivity and task complexity at the production-line level instead of the individual-level.[23] Results are presented in Table A4. They are less precise since they include all workers on the line, not just treated workers, but are very consistent with the individual-level results. The treatment effects for both efficiency and SAM are statistically significant at the 10% level after treatment. The magnitude of the line-level treatment effect for efficiency is about 40% of the direct treatment effect, and for SAM is about 70% the direct treatment effect. These results provide further evidence that the main results are not driven by differential attrition rates by treatment. Furthermore, they indicate that the firm gains not only higher individual-level productivity from training the treated workers, but that these workers enable the entire production lines on which they produce to become more productive.

6.6 Return on investment calculations

Figure 9: Total Program Costs Over Time



To quantify the total returns in terms of profit to the firm, we combine our treatment effect estimates on retention (person-days) and productivity with costing data obtained from the program administrators. We report in Table 7 calculations of the net present value of costs and benefits. Benefits are calculated in terms of additional person days and incremental productivity from treated workers using estimates from the randomized evaluation. Cost involve fixed and variable programmatic costs, lost productivity due to training, and wage increases (we do not report wage as a separate category of cost in Table 7 because these impacts were essentially negligible).[24] We omit spillover impacts from

[23]Note that these results include all workers on the production line, not just those that signed up for the program.

[24]In addition, we implicitly assume in calculating lost productivity due to reduced person days that the rate of hiring or

the calculations that follow to produce conservative estimates, given that the effects on productivity are not statistically significant.

Figure 10: Cumulative Program Benefits Over Time

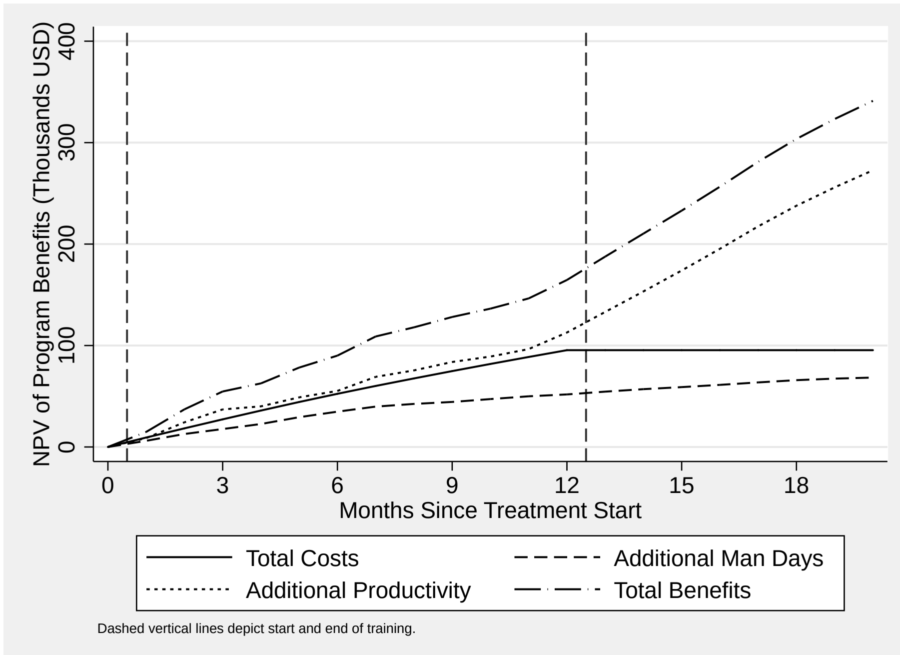


Table 7 first outlines costs of the program, both overhead costs and variable costs. The overhead costs are given by the costs of hiring two full-time trainers per factory for the 12 months of the program, additional support time from HR personnel, printed materials, food, and equipment (e.g., PA system). The variable costs are from lost production hours, and the marginal increase in wages for treated workers. For the 1087 treated workers, total program costs are approximately \$95,000, about \$57,000 of which are overhead costs, and the remainder variable costs. The time path of total costs in net present value (NPV) is shown in Figure 9, with total costs rising linearly during the program period and peaking at program completion.

Details on profit margins on additional revenue both from an additional person day and additional productivity, as well as additional revenue per garment were obtained from the firm. The benefits of the program are generated by the higher number of cumulative person days accrued to the firm and by higher worker productivity. At the end of the program period, the NPV of these benefits is just over \$164,000, about \$52,000 of which is the result of additional person days gained during the program and the rest due to productivity gains. At the end of our tracking period (8 months after program completion), total benefits are substantially higher, more than \$341,000. In the post-program period, returns via productivity gains dominate, accounting for more than 70% of the total benefits. Figure 10 plots the time path of cumulative benefits to the firm. Note that these returns accrued net of attrition

worker replacement is common across treatment and control lines such that differential attrition produces truly lost person days. This is largely true as hiring is centralized for each factory unit. Accordingly, firm management reported to us that it is impossible for the rate of recruitment, hiring, and training to respond to differential turnover across lines within factory unit.

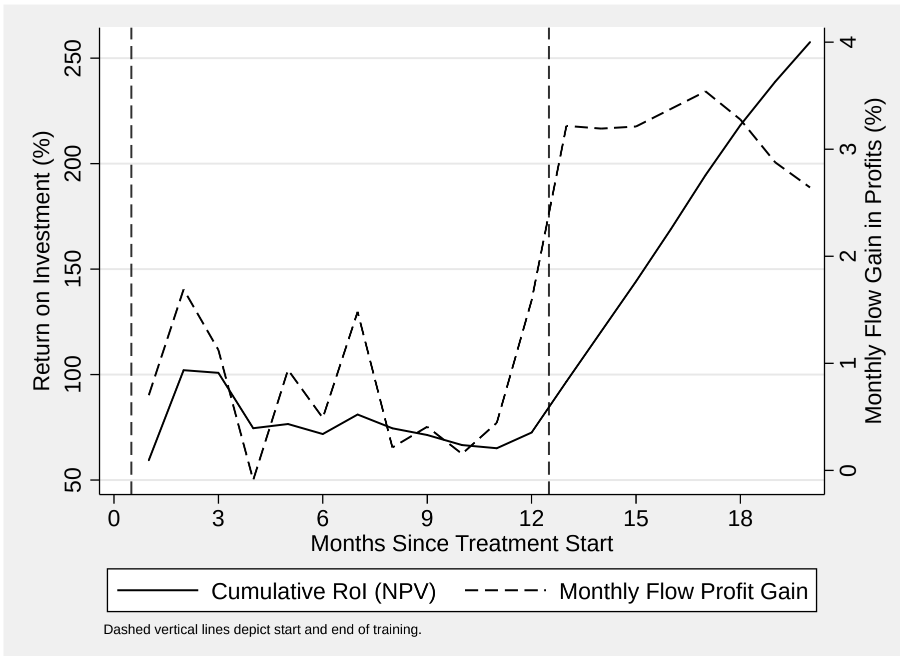
Table 7: Return on Investment Calculations (Costs and Benefits to Firm)

Sewing Department Only (1087 Treated Workers)	
P.A.C.E. Training Overhead Cost (Trainers, HR Oversight, Materials, and Food for 12 Mos)	-$57,091.68
P.A.C.E. Training Variable Cost (Lost Garments from Lost Man Hours)	-$38,314.88
Total Cost *(All numbers in present value)*	-$95,406.56
1 Year After Program Announcement	
Additional Person Days	$51,804.37
Additional Productivity	$112,785.00
Net Present Value of Subtotal	$164,589.30
Net Rate of Return	**73%**
20 Mos After Program Announcement	
Additional Person Days (End of Observation)	$68,389.79
Additional Productivity (Garments per 8 hr day)	$272,767.00
Net Present Value of Subtotal	$341,156.80
Net Rate of Return	**258%**
Assumptions	
Additional Garments per Additional Man Day	8.2
Additional Revenue per Garment	$7.00
Labor Contribution to Cost ("Cut to Make")	25%
Profit Margin on Additional Revenue from Additional Productivity	18.75%
Profit Margin on Additional Revenue from Additional Man Day	5%
Interest Rate	10%
INR per 1 USD	58

Notes: Trainer salaries were 17,000 INR per month for each trainer. There were 2 trainers for each of the 5 factories; 10 trainers in total. Additional HR personnel time for program oversight amounted to 6,659 INR per month across all 5 factories. Materials and equipment costs amounted to 26689 INR per month across all 5 factories, and food costs amounted to 27,175 INR per month across all 5 factories. Additional garments per additional man day is calculated by dividing the average worker level SAM (minutes to complete the operation on a single garment) by the line level SAM (minutes to complete a full garment for the line) and multiplying by 480 minutes in a work day. All additional productivity and man days coefficients are taken from the montly impacts estimated in the main results and appropriately scaled by the number original sample workers remaining in the factory in each month. Additional revenue per garment is taken from the accounting department of the firm, as is the "Cut to Make" or labor percent contribution to total production cost. Profit margin on additional revenue generated through improved efficiency is calculated as 75% of the "Cut to Make" cost as instructed by the accounting office of the firm and the profit margin on additional revenue from an additional man day is equivalent to the average profit margin of the firm. The monthly interest rate is the average interest rate that prevailed during the study time period. Similarly, the exchange rate is the average from the study period.

– that is, we only count person days gained and productivity increases accruing to workers who were still present at each point in time.

Figure 11: Cumulative and Flow Return Over Time



The net rate of return at the end of the program period is thus 73% (i.e., at program end, costs had been entirely recouped by the firm, plus 73 percent additional returns). Twenty months after program completion, flow benefits mostly from post-program productivity impacts help generate a net rate of return of 258%. Figure 11 shows the time path of the cumulative and flow net rate of return.

7 Conclusion

In this paper we study the labor market impacts of soft skills. We combine randomized placement into an on-the-job soft skills training program for female garment workers in India with detailed measurement of productivity, retention, wages, and other workplace outcomes, to characterize the effects ofthis training on workers as well as on the firm. We find that soft skills improvements generate large and persistent productivity impacts, but have negligible effects on wages and turnover. These results are consistent with theories of labor market imperfections, and suggest that the firm captures most of the gains from the increased marginal productivity of labor.

Growing interest in active labor market policies (Heckman et al., 1999) in low-income countries has spurred study of the impacts of vocational training programs, which often include a soft skills training component (Betcherman et al., 2004). In general, estimates of the labor market benefits of training alone (as opposed to training plus asset or cash transfers) do not yield consistent evidence of impact (McKenzie, 2017). Interventions focused on young women may be one area of exception – see, e.g., recent work by Buvinić and Furst-Nichols (2016) and Acevedo et al. (2017). This recent work, along with our findings, may pave the way for greater concentration on active labor market interventions focused on women workers.

Finally, our work is relevant to the literature on female labor force participation (LFP) and employment outcomes, particularly in low-income country contexts (Heath and Jayachandran, 2016). This policy question of how to increase the LFP and career growth of women is especially salient in India, where the level of female LFP is not only unusually low considering India's level of development (India ranks 120th out of 131 countries in female LFP (Chatterjee et al., 2015)), but has substantially decreased in rural areas between 1987 and 2009, despite a fertility transition and relatively robust economic growth (Afridi et al., 2016). Studying improvements in career prospects for women, via managerial training and promotion as Macchiavello et al. (2015) do, or via soft-skills training and resulting productivity enhancements as we do, can contribute to our understanding of determinants of female labor force participation that are amenable to policy intervention.

References

Acemoglu, D. (1997). Training and innovation in an imperfect labour market. *The Review of Economic Studies*, 64(3):445–464.

Acemoglu, D. and Pischke, J.-S. (1998). Why do firms train? theory and evidence. *The Quarterly Journal of Economics*, 113(1):79–119.

Acemoglu, D. and Pischke, J.-S. (1999). Beyond becker: training in imperfect labour markets. *The Economic Journal*, 109(453):112–142.

Acevedo, P., Cruces, G., Gertler, P., and Martinez, S. (2017). Living up to expectations: How job training made women better off and men worse off. Technical report, National Bureau of Economic Research.

Afridi, F., Dinkelman, T., and Mahajan, K. (2016). Why are fewer married women joining the work force in india? a decomposition analysis over two decades.

Aizer, A. and Cunha, F. (2012). The production of child human capital: Endowments, investments and fertility. *Unpublished paper, Brown University*.

Altonji, J. G. and Spletzer, J. R. (1991). Worker characteristics, job characteristics, and the receipt of on-the-job training. *Industrial & Labor Relations Review*, 45(1):58–79.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, pages 1481–1495.

Ashraf, N., Low, C., and McGinn, K. (2017). Negotiating a better future: Communication skills and inter-generational investment in zambia. Technical report, mimeo.

Attanasio, O. P., Fernández, C., Fitzsimons, E. O., Grantham-McGregor, S. M., Meghir, C., and Rubio-Codina, M. (2014). Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in colombia: cluster randomized controlled trial. *BMJ*, 349:g5785.

Autor, D. H. (2001). Why do temporary help firms provide free general skills training? *Quarterly Journal of Economics*, pages 1409–1448.

Barrett, A. and O'Connell, P. J. (2001). Does training generally work? the returns to in-company training. *Industrial & labor relations review*, 54(3):647–662.

Barron, J. M., Berger, M. C., and Black, D. A. (1999). Do workers pay for on-the-job training? *Journal of Human Resources*, pages 235–252.

Bartel, A. P. and Sicherman, N. (1998). Technological change and the skill acquisition of young workers. *Journal of Labor Economics*, 16(4):718–55.

Bassanini, A., Booth, A., Brunello, G., De Paola, M., Leuven, E., et al. (2007). Workplace training in europe. *Education and Training in Europe*.

Bassi, V., Nansamba, A., and Liberia, B. (2017). Information frictions in the labor market: Evidence from a field experiment in uganda. Technical report, mimeo.

Becker, G. S. (1964). Human capital.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, pages 289–300.

Betcherman, G., Dar, A., and Olivas, K. (2004). *Impacts of active labor market programs: New evidence from evaluations with particular attention to developing and transition countries*. Social Protection, World Bank.

Blundell, R., Dearden, L., Meghir, C., and Sianesi, B. (1999). Human capital investment: the returns from education and training to the individual, the firm and the economy. *Fiscal studies*, 20(1):1–23.

Borghans, L., Duckworth, A. L., Heckman, J. J., and Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of human Resources*, 43(4):972–1059.

Buvinić, M. and Furst-Nichols, R. (2016). Promoting women's economic empowerment: what works? *The World Bank Research Observer*, 31(1):59–101.

Campos, F., Frese, M., Goldstein, M., Iacovone, L., Johnson, H. C., McKenzie, D., and Mensmann, M. (2017). Teaching personal initiative beats traditional training in boosting small business in west africa. *Science*, 357(6357):1287–1290.

Chang, C. and Wang, Y. (1996). Human capital investment under asymmetric information: The pigovian conjecture revisited. *Journal of Labor Economics*, pages 505–519.

Chatterjee, U., Murgai, R., and Rama, M. (2015). Job opportunities along the rural-urban gradation and female labor force participation in india. *World Bank Policy Research Working Paper*, (7412).

Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.

Cunningham, W. V. and Villaseñor, P. (2016). Employer voices, employer demands, and implications for public skills development policy connecting the labor and education sectors. *The World Bank Research Observer*, 31(1):102–134.

Dearden, L., Reed, H., and Van Reenen, J. (2006). The impact of training on productivity and wages: Evidence from british panel data. *Oxford bulletin of economics and statistics*, 68(4):397–421.

DellaVigna, S., List, J. A., Malmendier, U., and Rao, G. (2016). Estimating social preferences and gift exchange at work. Technical report, National Bureau of Economic Research.

Deming, D. J. (2015). The growing importance of social skills in the labor market. Technical report, National Bureau of Economic Research.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S. M., and Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in jamaica. *Science*, 344(6187):998–1001.

Goux, D. and Maurin, E. (2000). Returns to firm-provided training: evidence from french worker–firm matched data. *Labour economics*, 7(1):1–19.

Grantham-McGregor, S. M., Powell, C. A., Walker, S. P., and Himes, J. H. (1991). Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the jamaican study. *The Lancet*, 338(8758):1–5.

Groh, M., Krishnan, N., McKenzie, D. J., and Vishwanath, T. (2012). Soft skills or hard cash? the impact of training and wage subsidy programs on female youth employment in jordan. *The Impact of Training and Wage Subsidy Programs on Female Youth Employment in Jordan (July 1, 2012). World Bank Policy Research Working Paper*, (6141).

Groh, M., McKenzie, D., and Vishwanath, T. (2015). Reducing information asymmetries in the youth labor market of jordan with psychometrics and skill based tests. *The World Bank Economic Review*, 29(suppl 1):S106–S117.

Guerra, N., Modecki, K., and Cunningham, W. (2014). Developing social-emotional skills for the labor market: The practice model. *World Bank Policy Research Working Paper*, (7123).

Heath, R. and Jayachandran, S. (2016). The causes and consequences of increased female education and labor force participation in developing countries. Technical report, National Bureau of Economic Research.

Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour economics*, 19(4):451–464.

Heckman, J. J., LaLonde, R. J., and Smith, J. A. (1999). The economics and econometrics of active labor market programs. *Handbook of labor economics*, 3:1865–2097.

Heckman, J. J. and Mosso, S. (2014). The economics of human development and social mobility. Technical report, National Bureau of Economic Research.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Technical report, National Bureau of Economic Research.

Ibarrarán, P., Kluve, J., Ripani, L., and Rosas, D. (2015). Experimental evidence on the long-term impacts of a youth training program.

Katz, E. and Ziderman, A. (1990). Investment in general training: The role of information and labour mobility. *The economic journal*, 100(403):1147–1158.

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S.-L. T., Manderscheid, R. W., Walters, E. E., et al. (2003). Screening for serious mental illness in the general population. *Archives of general psychiatry*, 60(2):184–189.

Konings, J. and Vanormelingen, S. (2015). The impact of training on productivity and wages: firm-level evidence. *Review of Economics and Statistics*, 97(2):485–497.

Leuven, E. and Oosterbeek, H. (2008). An alternative approach to estimate the wage returns to private-sector training. *Journal of applied econometrics*, 23(4):423–434.

Macchiavello, R., Menzel, A., Rabbani, A., Woodruff, C., et al. (2015). Challenges of change: An experiment training women to manage in the bangladeshi garment sector. Technical report, University of Warwick, Department of Economics.

McKenzie, D. (2017). How effective are active labor market policies in developing countries? a critical review of recent evidence. *The World Bank Research Observer*, 32(2):127–154.

Menzel, A. (2015). Organizational learning: Experimental evidence from bangladeshi garment factories.

Mincer, J. (1962). On-the-job training: Costs, returns, and some implications. *The Journal of Political Economy*, pages 50–79.

Montalvao, J., Frese, M., Goldstein, M. P., and Kilic, T. (2017). Soft skills for hard constraints: evidence from high-achieving female farmers.

Myer, L., Stein, D. J., Grimsrud, A., Seedat, S., and Williams, D. R. (2008). Social determinants of psychological distress in a nationally-representative sample of south african adults. *Social science & medicine*, 66(8):1828–1840.

Ng, J. (2013). Elicited risk and time preferences: the role of demographics, cognition, and interviewers. *University of Southern California Job Market Paper*.

Schoar, A. (2014). The importance of being nice: Supervisory skill training in the cambodian garment industry. Technical report, working paper.

Staritz, C. (2010). *Making the Cut?: Low-income Countries and the Global Clothing Value Chain in a Post-quota and Post-crisis World*. World Bank Publications.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

World Bank, D. P. G. (2012). *World Development Indicators 2012*. World Bank Publications.